Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **SEP 2007**	2. REPORT TYPE	3. DATES COVERED **00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE **An Examination of the Relationship Between Usage and Operating and Support Costs for Air Force Aircraft**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Rand Corporation,Pardee RAND Graduate School ,1776 Main Street,Santa Monica,CA,90407-2138**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**115**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

DISSERTATION

An Examination of the Relationship Between Usage and Operating and Support Costs for Air Force Aircraft

Eric J. Unger

This document was submitted as a dissertation in September 2007 in partial fulfillment of the requirements of the doctoral degree in public policy analysis at the Pardee RAND Graduate School. The faculty committee that supervised and approved the dissertation consisted of Edward Keating (Chair), Bart Bennett, and Lara Schmidt.

RAND PARDEE RAND GRADUATE SCHOOL

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

The Pardee RAND Graduate School dissertation series reproduces dissertations that have been approved by the student's dissertation committee.

Published 2008 by the RAND Corporation
1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213
RAND URL: http://www.rand.org
To order RAND documents or to obtain additional information, contact
Distribution Services: Telephone: (310) 451-7002;
Fax: (310) 451-6915; Email: order@rand.org

Abstract

This research examines the relationship between operating and support (O&S) costs and usage of Air Force aircraft, in order to improve resource allocation. Currently the Air Force uses an average cost metric to forecast costs related to flying hours. Problems arise with the accuracy of the cost per flying hour (CPFH) factors when the relationship between cost and usage is either nonlinear or includes nontrivial fixed costs.

Superficially, it may seem reasonable that if the Air Force flies an aircraft twice as many hours, O&S costs should double. However, empirical evidence shows that the doubling of flying hours actually increases non-fuel operating and support costs by less than that amount. This finding is consistent with nontrivial fixed costs, challenging the validity of the current proportional budgeting metric.

Another aspect to forecasting Air Force budgets is whether O&S costs vary with flying hours or with the number of aircraft. The Air Force currently groups O&S budget components into three cost categories: variable with flying hours, variable with the number of aircraft, and fixed costs. We find that the high correlation between flying hours and the number of aircraft prevents one variable from outperforming the other in predictive models. Fuel cost is the only category with clear statistical evidence to support the use of flying hours over aircraft inventory in predictive models.

The Air Force can improve its allocation of O&S resources by altering the current proportional CPFH metrics to better accommodate fixed costs. When we apply the findings of this research to budget projections, we see substantial differences to forecasts created with the status quo metrics. We compared estimates of the largest O&S component—the $11.5 billion budget for consumable parts and fuel. For a 10% increase in flying hours, our method estimates a budget approximately $406 million less than the proportional method. For a 10% decrease in flying hours, our method retains over $402 million more than the proportional method. This is an important difference for Air Force budget planning, particularly during the transitions periods that preceded and follow major contingency operations.

Table of Contents

Acknowledgments

Over the past fifteen years of my Air Force career, I have learned that no major achievement can be realized without the help, guidance, and encouragement of those around me. From completing my testing requirements for ACSC to defending program office estimates to the GAO, I have been the beneficiary of the aid and counsel of many outstanding individuals. However, my RAND experience is certainly unique in terms the magnitude and diversity of support that I received. I am humbled by the graciousness, professionalism, and intellectualism of my colleagues and mentors.

First and foremost, I am grateful for the counseling and patience of my committee, Ed Keating, Bart Bennett, and Lara Schmidt. Ed Keating, my chair, provided the proper mix of insight and positive (not normative) questions to keep me on the right course. He also holds the distinction of turning around comments quicker than any reviewer in the history of dissertations; I was amazed and impressed. Bart Bennett emphasized the importance of getting my writing started early. That was worthy advice, since he helped me expand the scope of my research with deeply intriguing ideas. Lara Schmidt was always the voice of reason in light of the statistics I intended to invent. Her explanations and comments were always clear, accurate, and helpful. I give special thanks to Dr. Michael Alles for his spirited and highly intellectual comments on my work.

There are numerous individuals at RAND who help me accomplish my goals. Natalie Crawford, the consummate mentor, helped me with issues ranging from TDY funding to career counseling. Hers was the first voice that welcomed me to RAND—a moment I will never forget. I would also like to thank Michael Kennedy, not only for the funding that allowed me to travel for data collection, but for his time. He and Fred Timson provided valuable input that helped turn the project around at a critical moment.

I would like to I would like acknowledge the stellar support I received from the Air Force. Tom Lies, Larry Klapper, Gary McNutt, and Crash Lively provided access to corporate knowledge that helped me understand the intricacies of AFTOC and O&S funding. Eric Hawkes and Patrick Armstrong, both AFIT alums, injected enthusiasm along with some lit review and data help—a very good combination. Although I was unable to use PDMSS data for this research, I would like to thank Mark Armstrong for his incredibly professional support of my inquiries.

The cornerstone of my PRGS experience was the interaction I had with my very accomplished colleagues. I would like to thank Ying Liu, Yang Lu, and Nailing "Claire" Xia for their help in teaching me Mandarin. Although I only learned two words, I use them both regularly in conversation. I must note that Claire and Yang showed great patience in explaining the nuances econometric theory to me. However, only Ying had the necessary endurance to discuss microeconomics. Ryan Keefe and Jordan Fischbach set the record at five for attending my dissertation briefings.

No one has had greater faith in my abilities than my parents. During my darkest moments, they held steadfast to the belief that I could finish this degree. I do not believe that I could have sustained the continuous effort required without their conviction. I also thank Angela for her support of my quest.

Acronyms

ABIDES	Automated Budget Interactive Data Environment System
AF	Air Force
AFI	Air Force Instruction
AFCAA	Air Force Cost Analysis Agency
AFCAIG	Air Force Cost Analysis Improvement Group
AFKS	Air Force Knowledge Services
AFTOC	Air Force Total Ownership Costs
ASD	Average Sortie Duration
AVFUEL	Aviation Fuel
CI	Confidence Interval
CONS	Consumables
CPFH	Cost per Flying Hour
CPI	Consumer Price Index
CY06	Constant Year 2006
DLR	Depot Level Reparable
DoD	Department of Defense
EEIC	Element of Expense and Investment Code
FHs	Flying hours
FYDP	Future Years Defense Plan
FY06	Fiscal Year 2006
GAO	Government Accounting Office
GPD	Gross Domestic Product
GPC	Government Purchase Card
GSD	General Support Division
MAJCOM	Major Command
MD	Mission Design
MDS	Mission Design Series
MILPERS	Military Personnel
MSD	Mission Support Division
O&M	Operations and Maintenance
O&S	Operating and Support
OLS	Ordinary Least Squares
OSD CAIG	Office of the Secretary of Defense Cost Analysis Improvement Group
POL	Petroleum, Oil, and Lubricants
POM	Program Objectives Memorandum
RECCE	Reconnaissance Aircraft
REMIS	Reliability and Maintenance Information System
SAF/FM	Secretary of the Air Force financial Management
TAI	Total Aircraft Inventory
VAMOSC	Visibility and Management of Operating and Support Costs

Figures

Tables

CHAPTER ONE

Introduction

This research improves the United States Air Force's (USAF) ability to manage the weapons system Operating and Support (O&S) budget by creating an alternative estimating strategy to the current top-level forecasting methods. Over the last ten years, O&S costs averaged about $40 billion annually. This research quantifies aircraft usage effects—the change in O&S costs induced by an additional unit of usage—controlling for the calendar costs of aging and aircraft variations. The identified usage effect improves methods of projecting future maintenance costs. With more accurate cost estimates, the Air Force will be better able to set and manage O&S budgets for its aircraft inventory.

For example, in a post-contingency environment, USAF may experience a drawdown in terms of both personnel and aircraft. It follows that these reductions, particularly reducing the number of aircraft, will affect average sortie durations, total annual flying hours (FHs), and annual FHs per aircraft. In this setting, a top-level analysis of the relationship between usage (flying hours) and O&S costs can provide improved estimates of future O&S budgets.

The relevant policy question is how to improve Air Force O&S resource allocation through better estimation methods. To investigate this policy question, the following research questions address specific areas that affect the Air Force's ability to estimate O&S costs:

- How do O&S costs vary with flying hours?

- Should we model O&S costs as a function of flying hours or the number or aircraft?

- How does the usage effect identified in this research impact Cost per Flying Hour (CPFH) metrics which the Air Force currently uses to determine O&S budgets?

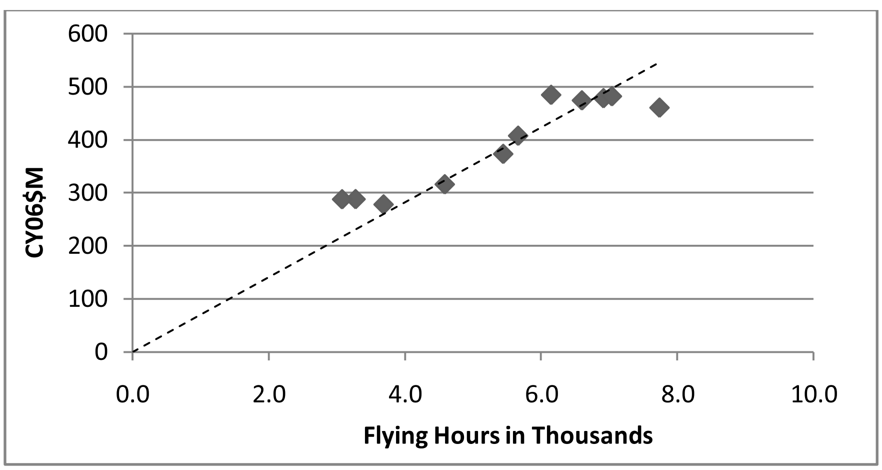
Establishing a quantitative relationship between aircraft usage and O&S costs hours will inform Air Force leadership on how to better allocate its budget resources. We provide an overview of issues related to our analysis, beginning with the limitations to the modeling strategy current employed by the Air Force.

Background

This study focuses on aircraft usage to better identify methods to predict future budgets. With the prospect of future reductions in flying hours, the Air Force will want to be particularly careful not to reduce its budgets too far—a possibility with current forecasting methods. More accurate allocation of budget resources will allow the Air Force's leadership to better plan force structure, flying hours, military personnel budgets and O&M budgets—all tightly controlled resources.

The Air Force uses two different estimating approaches to forecast O&S budgets: bottom-up and top-down. The AF combines estimates derived from both of these approaches to build final budgets. The bottom-up approach includes estimates of costs and flying hours that aggregate from individual wings (the highest-level organizational component of an Air Force base) to the Secretary of the Air Force level. The top-down approach creates cost estimates from aggregated historical data—a streamlined approach used to validate the other method. We focus on alternatives to the top-down approach to create an estimating strategy that complements current methods.

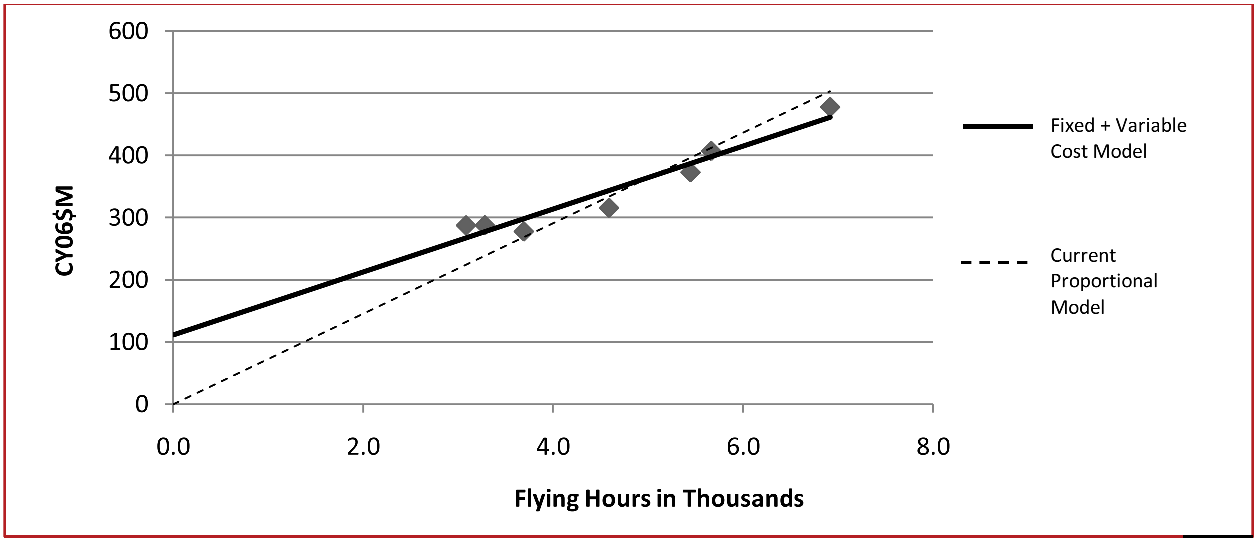
Figure 1-1 Current Proportional Model Example



At a top-level, the Air Force uses an average cost metric to predict costs, depicted by the dashed line in the Figure 1-1.[1] This average cost model is analogous to regression with the intercept forced to zero. The current metric implies that costs grow proportionally with flying hours; there are no costs associated with zero flying hours and as flying hours double, costs also double.

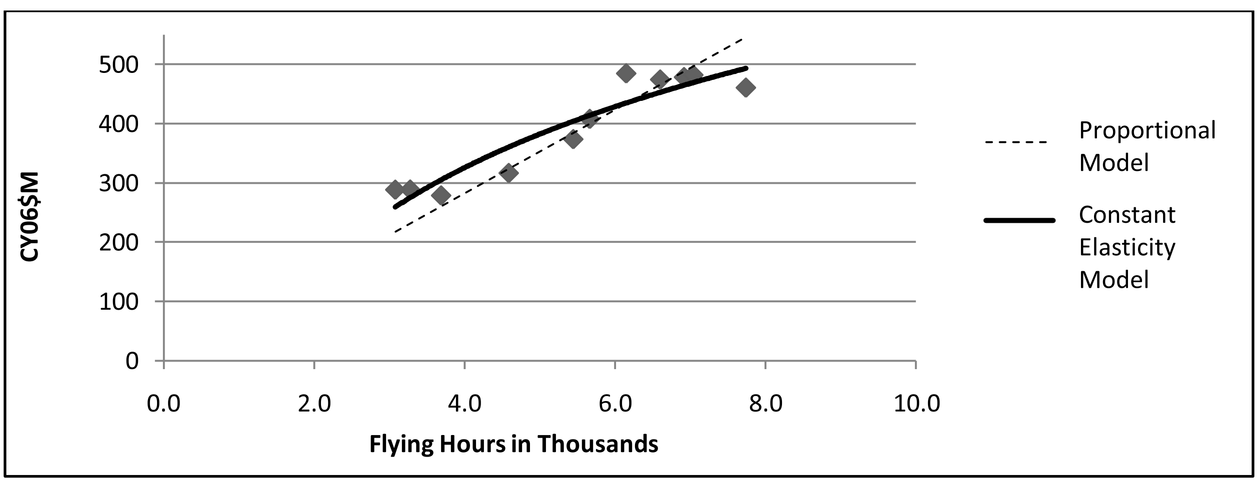
[1] The above scatter plot shows the flying hours and CY06 cost pairs for the B-2 from FY096 through FY06.

Figure 1-2 Proportional Model Compared to Fixed+ Variable Cost Model



However, two common situations can lead to incorrect estimates when using the proportional method. The first situation is the presence of fixed costs. As shown in Figure 1-2, nonzero fixed costs change the slope of the relationship between costs and flying hours, altering the estimate of costs within the relevant range of the data. This example shows that the presence of fixed costs tends to dampen the flying hour effect—a reduction or increase in flying hours from a given point will affect costs less than in the proportional model.

Figure 1-3 Proportional Model Compared to Constant Elasticity Model



Another potential problem with the proportional model is that it assumes a linear relationship between flying hours and costs. A nonlinear relationship between flying hours and costs can also alter estimates of cost within the relevant range. We show one possible nonlinear relationship in Figure 1-3. Here we depict the difference between the proportional model and a logarithmic estimate.

We assert that the Air Force can improve its O&S resource allocation by better accommodating fixed costs in their estimating strategies. The preferred method for estimating costs, depicted by the solid line in Figure 1-2, is to regress cost data on associated flying hour data by weapon system, allowing a non-zero intercept term—changing the current metrics to more accurately reflect fixed costs. Estimating costs for each system separately allows analysts to more easily control for unique aspects of each system.

However, the way the Air Force collects and reports cost data imposes a limit on the type of analysis that can be performed. While the preferred method discussed above would produce useful estimates of costs, there are an insufficient number of observations to estimate variable and fixed costs for each aircraft type. The Air Force routinely estimates costs with ten or fewer observations, causing nontrivial problems with the underlying assumptions for regression modeling.

We take an alternative estimating approach to overcome the sample size problems with system-specific estimations. We relate log-transformed costs to log-transformed usage variables, including observations for 34 aircraft platforms. Transforming into logs mitigates large variances in usage and costs between systems. This specification estimates a common elasticity with data from all included systems. We control for system in these top-level models to accommodate the differences between systems. Our approach assumes that there is a common flying hour–cost elasticity across platforms—that different aircraft platforms do not have fundamentally different relationships between flying hours and total costs. Our estimation strategy is intended as a complementary approach to the existing methods.

Can we identify a Usage effect?

Historically, it has been difficult to estimate usage effects on top-level costs since there has not been sufficient variation in flying hours—activity above or below baseline levels. However, operations in Iraq and Afghanistan provide a natural experiment—a demarcation in time where there are discernable changes in flying hours and flying hours per Total Aircraft Inventory (TAI). This recent marked variation in flying hours allows econometric models to discern whether or not usage has an effect on O&S costs and identify the character of that effect. In short, contingency operations allow us to address the primary research question: can we use heterogeneity in FHs to estimate the usage effect on O&S costs?

Figure 1-4 Flying Hour Trends by Fiscal Year

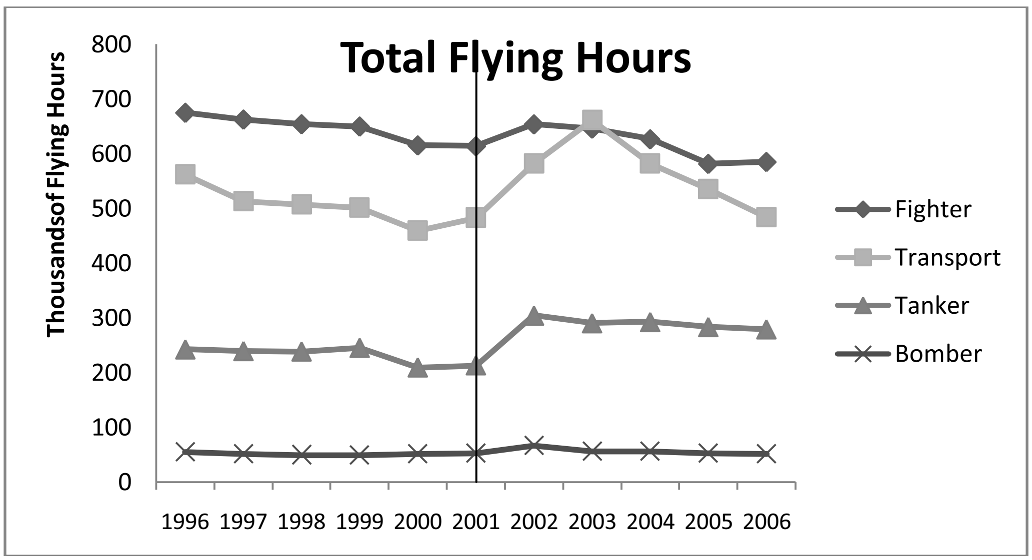
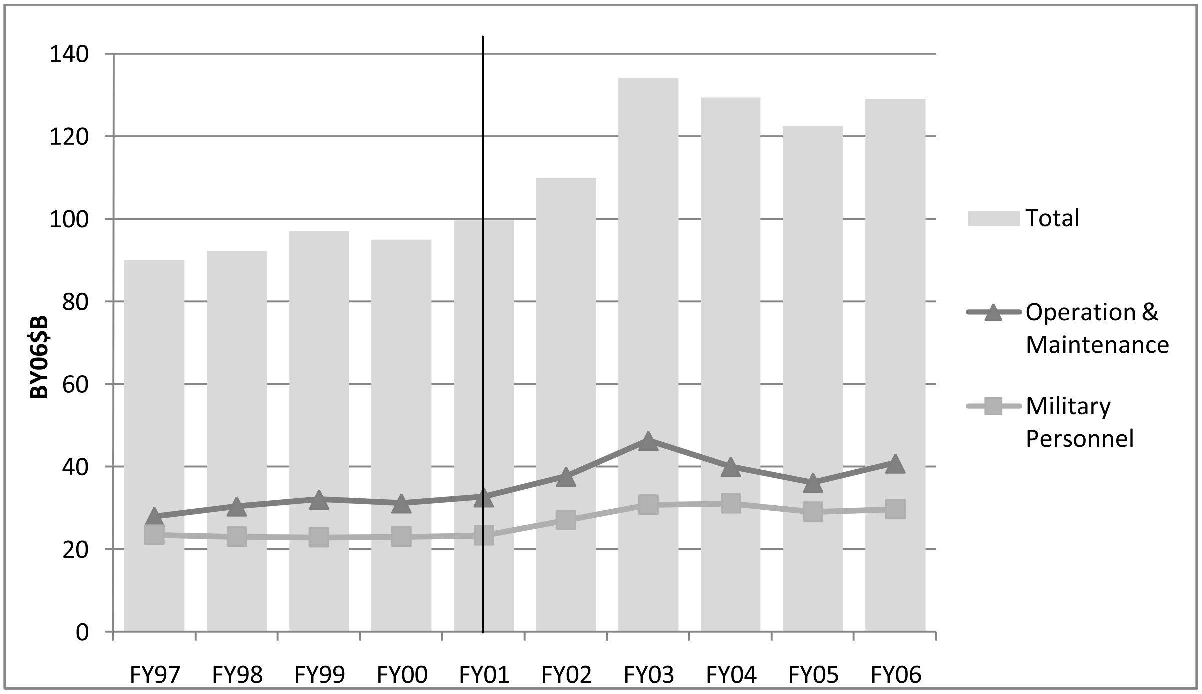


Figure 1-4 shows that contingencies in Iraq and Afghanistan have dramatically increased flying hours, particularly for Tankers and Transports. There is a similar effect for reconnaissance (RECCE) and Special (not shown), but it would be suppressed by the scale of the chart. The vertical line in Figure 1-4 and subsequent figures indicates the last year before contingency operations began in Afghanistan and Iraq; we include the line to highlight the changes in flying hours or budget dollars from peacetime operations to wartime operations. The variability in total flying hours provides an opportunity to identify a relationship between usage and costs, if one exists.

The increase in flying hours was associated with an impact on the Air Force total budget. Figure 1-5 depicts USAF budget trends by appropriation from fiscal year 1997 (FY97) to fiscal year 2006 (FY06) in base year 2006 terms—essentially real 2006 dollars (Air Force Magazine, 2006). The budget figures shown are from the Budget Authority as presented in the President's Budget.

Figure 1-5 USAF Budget in BY06$B by Appropriation



Congress appropriates military funds by use; the O&M appropriation and military personnel appropriation are the two largest appropriations that fund Operating & Support functions. We show the total Air Force budget and these two primary O&S appropriations in Figure 1-5. In BY06$, the total AF budget increased from $93.7B in FY97 to $130.2B in FY06; over the same period O&M costs increased from $29.1B to $40.0B, while personnel costs increased from $24.6B to $30.4B . The figure also shows that the general trend for both of the appropriations is increasing, even controlling for the spike for contingency operations in FY03 and FY04. This indicates that the costs for each appropriation are growing in real terms. Note that O&M budgets are the largest single appropriation for the Air Force comprising over 25% of the total Air Force budget.[2]

Figures 1-4 and 1-5 suggest that there is a an association between the O&M and Military Personnel portions of Air Force budget and flying hours—as flying hours increase, costs increase. However, this is a correlation and not necessarily causation. Possible aircraft system O&S cost drivers include, usage,

[2] Air Force Statistical Digest, FY05 Data Tables

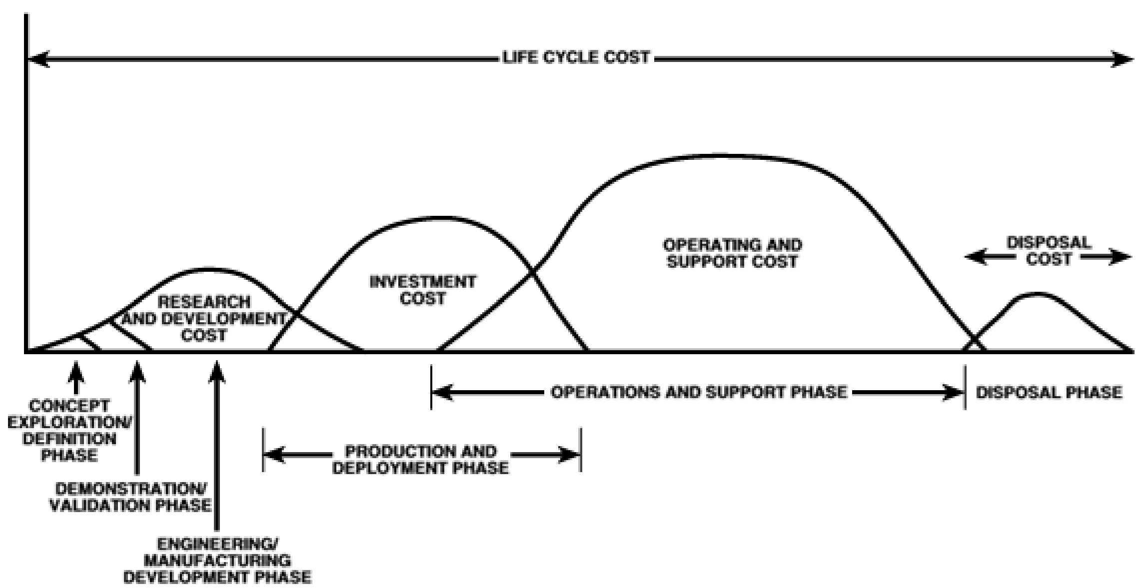
aging, number of aircraft, workforce reductions, depot closures, spare parts shortages, and a host of other issues (Greenfield and Persselin, 2002). The major analytical effort of this study focuses on usage effects, but age will be controlled for in the analysis. Chapters Four and Five will address the relationship between flying hours and O&S costs and Chapter Six will discuss the impact of the effect on the Cost per Flying Hour (CPFH) metric. In order to explore these relationships, we require further elaboration on the terminology and structure associated with AF cost and usage data.

Overview of O&S Cost Data

Weapon System Operating & Support (O&S) Costs

This report focuses on Operating and Support (O&S) costs for weapon systems—a sizeable subset of the total Air Force total O&S budget. Congress funds weapon system O&S activities through several different appropriations, including operations and maintenance (O&M) and military personnel (MILPERS) shown above in Figure 1-5. Weapon system O&S includes the costs associated with operating, maintaining, repairing, and supporting Department of Defense (DoD) weapons and other equipment, as well as pay and other benefits for military and civilian personnel.

Figure 1-6 Life Cycle Cost (OSD CAIG, 1992)



O&S costs comprise the majority of life cycle costs for most weapon systems. Aircraft systems typically spend over 60% of their total life cycle costs in O&S. For example, the Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG) reports that the F-16 system will spend about 78% of its life cycle costs in O&S (OSD CAIG, 1992). Figure 1-6 illustrates the typical life cycle of weapons system program and shows the allocation of costs to a given program phase. In the next section, we explain the DoD-accepted structure for O&S costs as established by OSD CAIG guidance.

Cost Analysis Improvement Group (CAIG) O&S Cost Elements

The Cost Analysis Improvement Group (CAIG) is Office of the Secretary of Defense (OSD) level organization charged with the responsibility to conduct independent cost assessments of major defense acquisition programs, serve as the advisor to milestone decision authorities on life cycle costs, and establish cost analysis procedures and policies. In this capacity, the OSD CAIG provided cost structure for O&S programs that is useful for both creating estimates and collecting cost information (DoD, 2006). The CAIG established this structure, in part, to ensure consistency of reporting between the military services.

Table 1-1 shows the CAIG Operating and Support (O&S) costs for fiscal year 2006 (FY06) in constant year 2006 dollars (CY06$). Constant year dollars control for inflation and allow for direct comparison between fiscal years. The CY06 costs shown in the table and used in the baseline set of models were calculated using OSD inflation indices. While the military produces most cost estimates using OSD rates to correct for inflation, there are other indices that might account for inflation more appropriately. For example, the Consumer Price Index (CPI) or Gross Domestic Product (GPD) deflator could be used to adjust for inflation. We address the impact of altering inflation rates in Chapter Four.

Table 1-1 Abridged CAIG O&S Cost Breakout (Billions $CY06)[3]

CAIG Element	FY06$B Costs
1.0 Mission Personnel	9.53
1.1 Operations Personnel	2.56
1.2 Maintenance Personnel	5.75
1.3 Other Mission Personnel	1.22
2.0 Unit-Level Consumption	11.30
2.1 Fuel/Energy Consumption	5.60
2.2 Consumables	1.03
2.3 Depot Level Reparables	3.86
2.4 Training Munitions	0.35
2.5 Other Unit-level Consumption	0.47
3.0 Intermediate Maintenance	0.01
4.0 Depot Maintenance	2.93
4.1 Aircraft Depot Maintenance	1.98
4.3 Engine Depot Maintenance	0.66
4.4 Other Depot Maintenance	0.29
5.0 Contractor Logistics Support	3.24
6.0 Sustaining Support	0.43
7.0 Indirect Support	3.03
7.1 Personnel Support	0.38
7.2 Installation Support	2.66

The information shown in Table 1-1 was accessed in the Air Force Total Ownership Cost (AFTOC) database, the Air Force's primary weapons system cost reporting system. Although AFTOC reports costs in a variety of formats, we use the OSD CAIG format. There are seven "level-one" CAIG categories, 1.0 Mission Personnel through 7.0 Indirect Support. The level-one categories are subdivided into 22 "level-two" categories (not all shown) that further segregate costs in the level-one categories. The models that we create using the first level breakout of costs—1.0 Mission Personnel, 2.0 Unit Level Consumption without 2.1 fuel , etc.—we term "level-one models." We will also create models at the lowest level shown in the chart. Models including 1.1 operations, 1.2 maintenance, etc., we term "level-two

[3] We exclude the breakout for CAIG element 5.0 since nearly all of its costs are in a single sub-element, 5.2. We exclude 6.0 sub-elements due to their small size. We include a complete table for CAIG level-two O&S Cost for FY96-FY06 in the Appendix.

models."[4] Table 1-1 shows a complete list of the level-one costs we use in our Chapter Four analysis of the relationship between flying hours and O&S costs. It also shows a partial list of the level-two O&S costs we will use in our Chapter Five analysis of fixed and variable costs.

Within the CAIG breakout of costs shown in the figure, AFTOC provides normalized, annual cost of ownership data by aircraft model—also referred to as mission design (MD). The costs in the above table represent costs for the 34 MDs included in the study, not the entire AF inventory. We use Constant Year 2006 gross obligations as our primary dependent variable in this study; we examine models at lower CAIG levels to determine the relationship of usage and specific cost areas. The following figure depicts historical trends for level-one CAIG elements.

Figure 1-7 CAIG Level-One Element Gross Obligations by Fiscal Year (CY06$B)

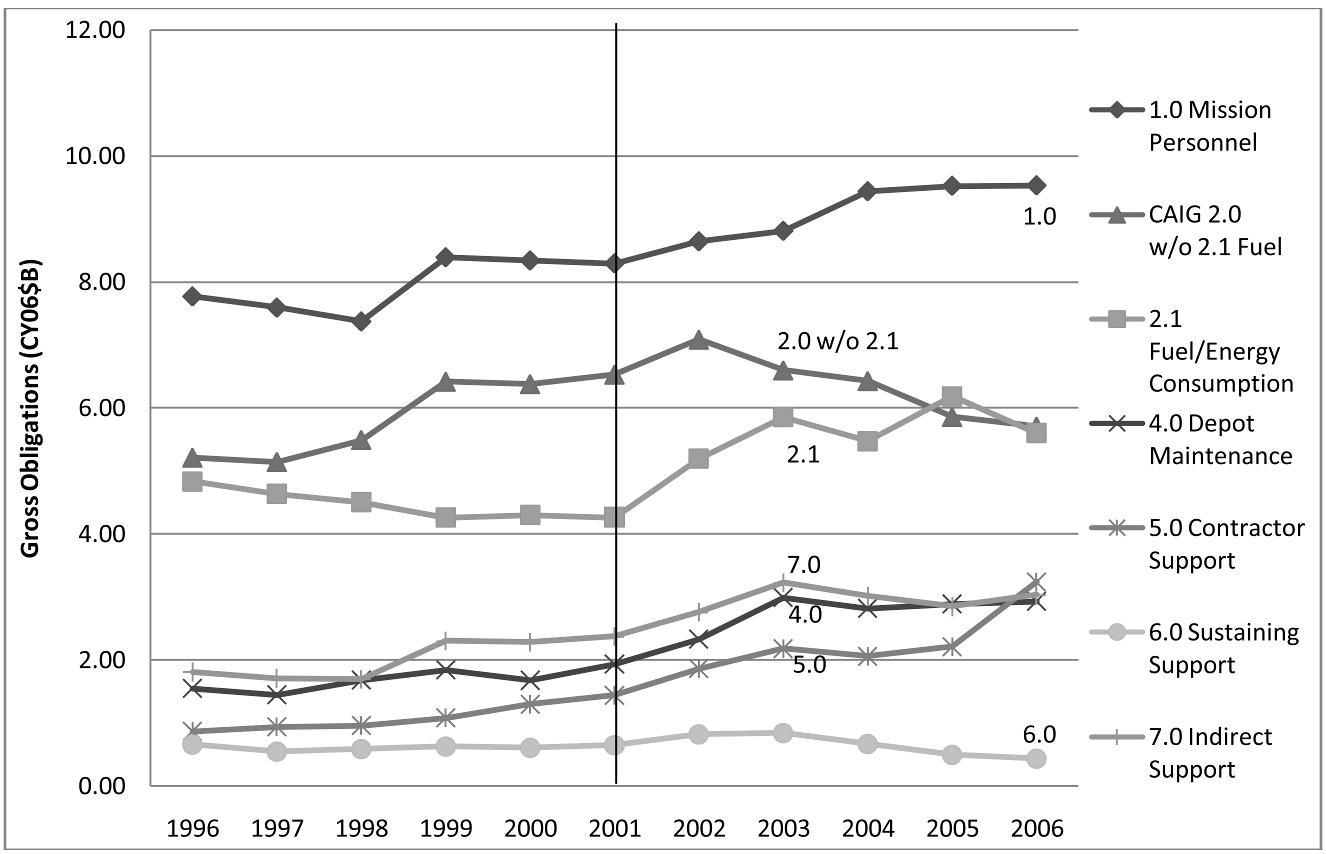


Figure 1-7 shows the 11-year trend for each of the level-one CAIG element costs in CY06$B, but excludes element 3.0 due to its extremely low relative cost. Element 1.0 Mission Personnel includes pay for

[4] The Appendix includes a more thorough description of the CAIG cost elements.

military and government personnel (for AF budgeting purposes, this is a combination of at least two separate appropriations). Element 1.0 comprises about 35% of the O&S budget, on average, though its percentage of total budget also has decreased slightly in recent years.

While Element 2.0 Unit level Consumption is the largest component of O&S costs, comprising about 40% of the O&S budget, we separate out the cost of fuel from all of the other costs in the 2.0 element since our aggregated cost models will not include fuel. The CAIG 2.0 without 2.1 fuel costs includes consumable maintenance material and depot-level reparable items, representing about 23% of the O&S budget on average. The cost for these non-fuel consumable items spiked in FY03 and has steadily decreased since. Fuel costs, element 2.1, have increased after a decline prior to FY02. While the amount of fuel consumed is related to flying hours, the cost associated with 2.1 shows the price volatility described earlier in FY04-FY06. During that period, flying hours decreased, but FY05 shows a marked increase in the budget for fuel. Element 5.0 contractor support has shown a marked increase since FY96, a possible result of the AF using a greater amount of contractor logistics support due to contingencies and acquisition decisions. The increase in CAIG 5.0, augments the increase in the 1.0 element—Mission Personnel.

FY02 is the first year where a substantial jump in costs is visible. However, the scale of the charts tends to suppress the jumps in smaller elements, such as 4.0, 5.0, and 7.0. FY03 saw continued growth in CY costs for all of the elements, except the non-fuel 2.0 element. So, while the charts may not indicate a significant jump in costs, model specifications that use just 4.0, 5.0 or 7.0 costs may show a dramatically different relationship with flying hours or Total Aircraft inventory (TAI) than the total cost relationship—emphasizing the need to classify appropriately cost dependencies on usage variables.

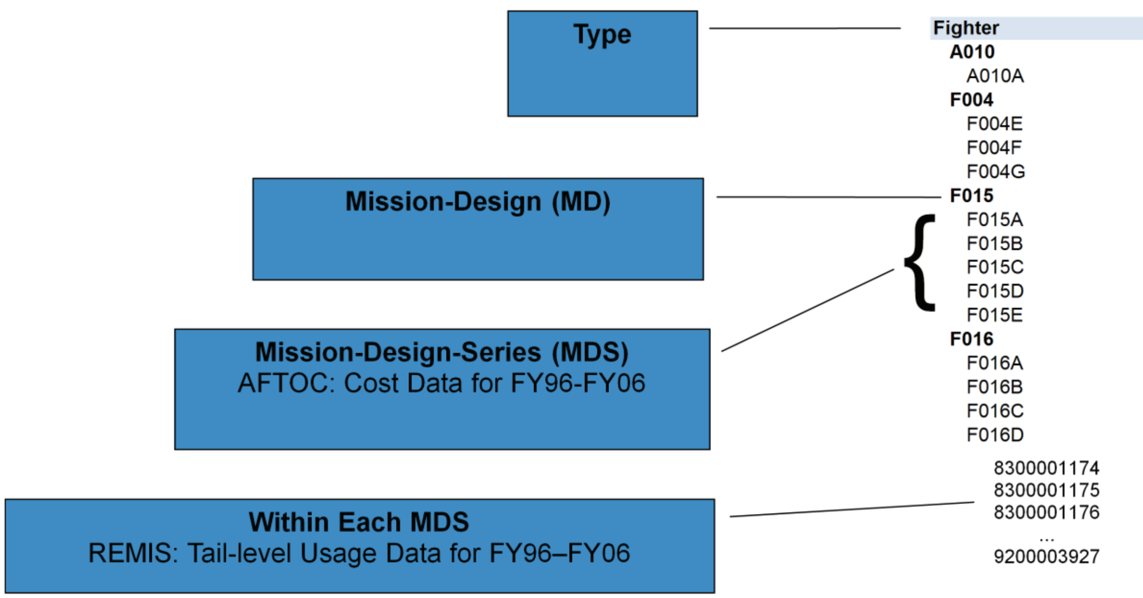
O&S costs are the dependent variable in all of our econometric models. We exclude Element 2.1 POL—Petroleum, Oil, and Lubricants (POL)—from all of the aggregated specifications, due to its known near-linear relationship with flying hours. POL on average accounts for nearly 15% of the total O&S budget and its proportional relationship with flying hours would bias our results toward the proportional model. For example, our cost elasticity estimate for the coefficient of flying hours is 0.56 for CAIG element 2.0 excluding 2.1; as flying hours double, costs increase by 56%. The elasticity estimate for CAIG 2.1 POL is 1.09; had we included 2.1 POL, our elasticity estimate for CAIG 2.0 would have been 0.888—far closer to the proportional model than with fuel excluded.

Air Force Aircraft Nomenclature and Usage Data

While the AFTOC system provides our dependent variable, we turn to another system for usage data. The Reliability & Maintainability Management Information System (REMIS) collects wide variety of aircraft data for multiple AF systems. In contrast to the higher level reporting of cost data in AFTOC, REMIS collects and reports information at individual aircraft or "tail" level. Our analysis explores the use of several "usage" variables, including flying hours (FHs), sorties, and landings. The Air Force groups usage data, as well as cost data, in a hierarchical structure.

Figure 1-8 shows the structure and nomenclature of USAF aircraft categorization. The highest level of aggregation is the aircraft type, which includes bombers, fighters, tankers, trainers, RECCE, special, and transport. Each type of aircraft comprises multiple Mission Designs (MDs). For example, the Fighters type contains MDs such as F-15 and F-16. Each MD, in turn, comprises one or more Mission Design Series (MDS). For the F-15 MD, there are five MDSs: F-15A, F-15B, F-15C, F-15D, and F-15E. Each individual airplane or "tail," can be categorized into a single MDS.

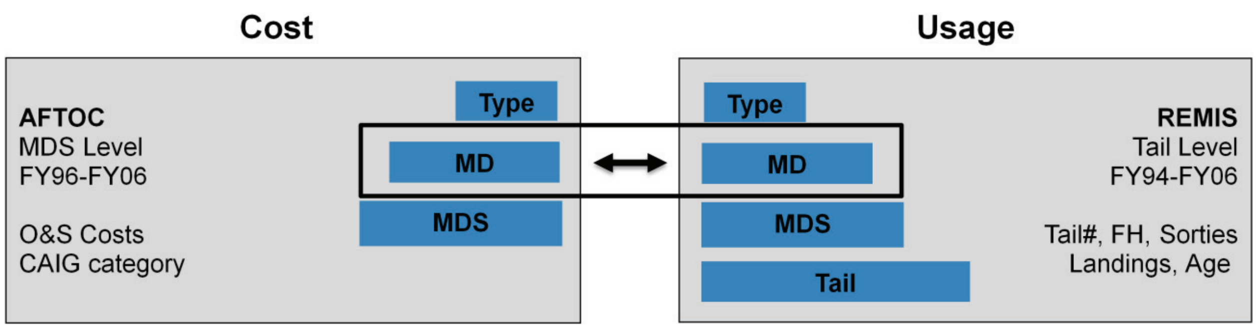
Figure 1-8 Nomenclature of USAF Aircraft



Cost Allocation Imposes Limit

While usage data are collected and reported for individual tails, the cost reporting systems cannot match this level of reporting fidelity. The allocation of budget obligation data from their original categories into the MD and MDS hierarchy poses significant problems to our analysis. The Air Force does not necessarily budget by MD, MDS, or tail. Rather it uses program budgets that may include multiple MDs or MDSs. From the program budgets, then, systems like AFTOC need to allocate obligation information to both MD/MDS structure and the CAIG structure.

Figure 1-9 Analysis Performed at the MD Level



A consequence of having to allocate budgets to the CAIG structure is that there may not always be relevant information with which to make the allocation. Typically, the AF collects costs at the MD level, but may lack the ability to apportion directly to the MDS level. In these cases, described further in Chapter Three, AFTOC allocates from MD to MDS using flying hours. Using flying hours to allocate costs biases estimates run at the MDS level; we mitigate the allocation problems by running models at MD level, as depicted in Figure 1-9. In short, we build MD level models to avoid overstating the cost relationship with Flying Hours.

A primary consideration in this study, both in methodological approach and eventual conclusions, is that Air Force O&S cost reporting imposes severe limits on analysis. These limits are present in two reporting dimensions: system aggregation level and frequency. In contrast to the airplane-specific information available for usage variable such as flying hours, AFTOC reports aggregated costs for a particular group of aircraft, not individual planes. Additionally, whereas most usage data are available monthly, AFTOC reports only annual cost data.[5] These restrictions result in having only one observation per year, per aircraft system. For the B-1 Lancer, then, we have access to on 11 cost-usage observations from FY96 through FY06. The sample size issue is the central problem with system-specific estimations; we attempt to overcome this limitation by estimating a common usage effect. The literature review in the next chapter addresses some of the top-level approaches researchers have used in the context of cost data limitations.

[5] AFTOC has implemented quarterly cost reporting as of Fiscal Year 2006.

CHAPTER TWO

Literature and Prior Work

This chapter discusses the prior work relating to Cost Per Flying Hour (CPFH) calculation and prediction using CPFH metrics. We include a limited discussion on aging aircraft literature, since the analysis techniques are similar, but focused on a different parameter of interest. Much of the work highlights problems with using flying hours to predict the parameters of cost and demand for spares.

Overview of Cost Per Flying Hour

The regular Air Force, Air National Guard, and Air Force Reserve, spend a combined $10 Billion a year on average for their flying hour programs, representing over 35% of the AF's annual Operating and Support (O&S) budget. This expenditure covers flying hour program requirements for consumable supplies, spare parts, and aviation fuel, essentially CAIG Element 2.0. [6]

In order to place the O&S usage effect in the proper context, we need to explain the Cost Per Flying Hour (CPFH) budget formulation process. The CPFH metric, referenced in AFI 65-503 AF Cost and Planning Factors, is the primary metric the Air Force uses to create future budgets from historical costs and flying hours. The creation of individual, MDS-specific CPFH factors is a multi-step process that involves many stages of input and review (Rose, 1997).

Figure 2-1 Simplified CPFH Metric

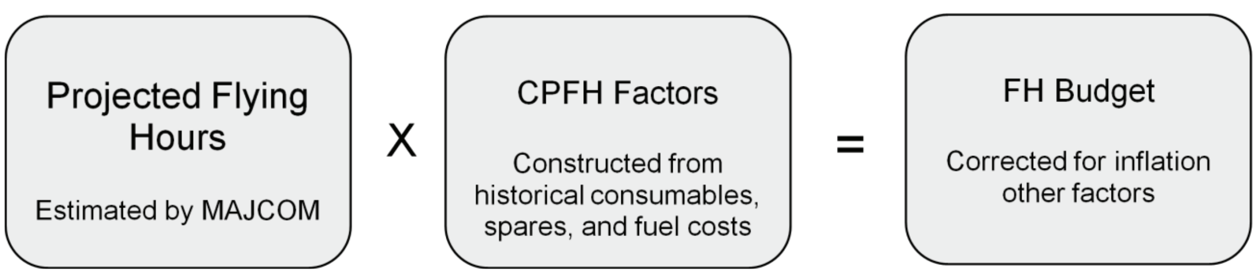


Figure 2-1 shows a simplified representation of the CPFH metric—the process by which the Air Force translates projected flying hours and MDS-specific CPFH factors into budgets. We include the full Air

[6] The costs associated with Cost Per Flying Hour pertain to CAIG Costs Element 2.0. However, the Air Force uses similar metrics to construct budgets for the other CAIG cost elements.

Force CPFH process diagram in the Appendix. Using input from the wings, MAJCOM[7] analysts create a CPFH factor for a given Mission Design Series (MDS), such as F-15E. The MAJCOM multiplies each of these factors by the projected flying hours to arrive at a flying hour budget. MAJCOM CPFH factor estimates are subject to reviews by numerous Air Force organizations, including the Air Force Cost Analysis Agency (AFCAA), before being used to construct budgets.

Projected Flying Hours

MAJCOMs forecast flying hour requirements based input from individual flying units within the command. Each unit within a given MAJCOM estimates the number of future sorties using a model with inputs that include the number of pilots, pilot experience, number of staff officers, training needs, and other activities. In estimating the number of sorties, these models assume that each unit will have all its aircraft and personnel assigned, creating an upper bound. Each unit converts the projected sorties to flying hours using unit-specific averages sortie durations (ASD). The average sortie duration varies among units according to geographic location, proximity to training ranges, and the type of aircraft the unit flies. The result of the calculation is the unit's flying hour requirement. The total flying hour requirement for all units in a particular command is provided to the command's financial management staff for use in developing the budget for flying hours. (GAO, 1999)

CPFH Factors and Budget Formulation Process

One of the key determinants of Air Force flying hour budgets is the CPFH factor. CPFH factors originate at the MAJCOM, but go through an extensive review before they are submitted to Air Force headquarters for approval; SAF/FM, AFCAA, and other Air Force organizations play significant roles in the CPFH factor development and validation process.

The CPFH factor is the ratio of recent historical O&M costs to their associated total flying hours. Each MAJCOM creates MDS-specific CPFH factors for spare parts, aviation fuel, and consumables. The Air Force budgets for spares, fuel, and consumables in four budget categories called Element of Expense and Investment Codes (EEICs). (AFCAIG, 1999) The Spare parts category includes EEIC 644: Reparable

[7] The Air Force is organized in a hierarchical structure. Major Command (MAJCOM), organized by geography and function, is the second highest category, below headquarters USAF. For additional information: http://www.maxwell.af.mil/au/afhra/rso/rso_index.html.

aircraft parts are those parts that can be repaired, usually by a depot, and are used in direct support of aircraft maintenance (e.g., aircraft engines). Aviation fuel includes EEIC 699: Aviation fuel and EEIC 693: Non-flying aviation fuel used for engine repair activities. The consumables category includes EEIC 609: aircraft parts that are not repaired—such as nuts and bolts—but are purchased through base supply and EEIC 61952: Consumable aircraft parts purchased outside base supply (Cooper, 2007). MAJCOM Analysts use five years of data to compute fuel requirements, with the other categories using two years of data.

The Air Force determines its flying hour budget—funding to support the daily operation of aircraft—from these CPFH factors and projected flying hours. For each MDS, analysts multiply CPFH factors by projected FHs to estimate an associated flying hour budget, as described in Figure 2-1. Once the approved MDS estimates are completed, the MAJCOM aggregates the budget requirements and submits them to the central Air Force budget system, called the Automated Budget Interactive Data Environment System (ABIDES). Air Force decision makers use requirement information in ABIDES to request budget authorization from congress. (AFCAIG, 1999)

Problems with the CPFH Metric

The construction of CPFH factors by dividing total costs by flying hours implies an average cost factor. Multiplying an average cost factor by projected flying hours to arrive at a budget forecast may misestimate budgets, given the presence of fixed costs. Fixed costs in the average cost factor would cause an exaggeration in estimated budget for a given number of flying hours. Budget forecasts for the flying hour program should use a marginal cost metric—fixed plus variable costs.

Another important aspect of the CPFH metric, as stated specifically in AF instruction 65-503, is that the Air Force uses these metrics for estimating both initial budgets and incremental changes within a fiscal year. From 65-503 Table A2-1 2006 Logistic Cost Factors, "Flying Hour Consumable Supplies (GSD) factors are used in the programming and budgeting process to build as well as increment and decrement consumable supply requirements based on changes in flying hours." This approach assumes that there are no fixed costs present in the procurement of GSD supplies and the possibility of volume discounts, for example, would disturb this structure. The 65-503 guidance notes for several of the logistic cost factors that the proportional metric should only be used for building or incrementing budgets, but not both. However, the proportional construction of the CPFH factors assures that there will be difficulties with both of these functions in the presence of fixed costs.

Since, projected flying hours frequently change due to weather considerations, real-world deployments, and other unanticipated events, the Air Force would encounter estimating problem routinely. For example, the GAO observed that from FY95 to FY98, the Air Force routinely flew about 90% of the flying hours for which they requested funding (GAO, 1999). Excess O&M funds were redirected to other O&M activities. The concern in a reduced flying hour scenario is that an average cost metric may decrement budgets too far. If the Air Force determines excess funds using the average CPFH metric, it risks underfunding the flying hour program.

The two primary concerns for using an average CPFH metric is that it exaggerates estimates in the presence of fixed costs and that it improperly adjusts budgets for within-fiscal year changes to flying hours. We will address these forecasting issues specifically in Chapter Six.

Cost per Flying Hour Research

We turn now to related CPFH research that examines methods to improve cost prediction. While some of the work analyzes top-level estimating strategies, similar to our analysis, many of the studies identify ways to improve system-specific estimation strategies.

Hildebrandt and Sze (1990) Found O&S Costs Increase Less Than Proportionally with Flying Hours

Hildebrandt and Sze (1990) construct regression models that relate flying hours to several different sub-elements of recorded O&S costs. Their study used data from the Weapon System Support Cost module of the Visibility and Management of Operating and Support Costs (VAMOSC) database.[8] Similar in approach to our study, Hildebrandt and Sze examine the relationship between VAMOSC cost data and flying hours per aircraft, the number of aircraft, flyaway costs, aircraft type, initial operational capability

[8] VAMOSC is a predecessor system to the Air Force Total Ownership Cost (AFTOC) system used in our analysis.

(IOC) year, and average aircraft type age. [9] They use a log-log regression specification, so the interpretation of the model coefficients is a constant elasticity.

The authors find that at the total O&S cost level, flyaway cost is an acceptable proxy for aircraft type and IOC year. In general, they also find that O&S costs increase less than proportionally with flying hours—as flying hours double, costs will less than double. For total O&S cost per aircraft, they find that a one percent increase in flying hour per aircraft results in a 0.62 percent increase in costs. Similarly for depot maintenance costs, they find that a one percent increase in flying hour per aircraft results in a 0.51 percent increase in costs. We extend their work by increasing the period of the dataset from six fiscal years to eleven, including more modern aircraft,[10] and analyzing costs at a finer resolution. Chapter Four addresses our analysis of the relationship between total O&S costs and flying hours.

Slay (1995) Showed Limitations of Flying Hour-Based Forecasting

Slay (1995) asserts that wartime changes in flying hour programs caused significant increases in the predicted spare parts costs that were not credible. Sizeable increases of flying hours caused large increases in predicted spare costs that were not realized. The Air Force stated that the problem derived from using flying hour-based forecasting methods developed using peacetime data; these methods overpredicted surge demand during wartime. The author wanted to determine a better way to forecast wartime spares—one that could use peacetime data more effectively.

Slay (1995) noted that one of the key differences between peacetime and wartime flying programs was that the average sortie durations increased appreciably in wartime. Another observation was that a method based on sorties would underestimate demand for spares. While flying hours increased during wartime, the number of sorties did not increase proportionally. In other words, the average sortie duration increased.

The solution was to create a model that included both flying hours and sorties, but that would allow for differences in relationships between the parameters and costs by different aircraft MDSs. Ultimately, the author suggested forecasting models where 90% of the costs derived from MDS sorties-based model and 10% of the costs from the MDS flying hour-based model. We tested Slay's specification that

[9] In this study, aircraft number refers to as Total Aircraft Inventory (TAI) and aircraft type refers to Mission Design (MD).
[10] The dataset in our research includes observations from FY96 through FY06. Hildebrandt and Sze (1990) include observations from FY81 through FY86.

included both sorties and flying hours as covariates. However, we found that the high degree of correlation between the two parameters prevented using them simultaneously.

Sherbrooke (1997) Found Average Sortie Duration Impacts Spares Forecasting

A variation on Slay's approach was to employ a model specification that included average sortie duration instead of sorties. Data from Operation Desert Storm[11] provided further evidence of Slay's observation that tactical aircraft flew longer sorties than during peacetime, causing an overstatement of demand for spares.

Sherbrooke(1997) refined Slay's work by presenting regression models for 24 MDSs that related spares demand to the sortie number of the day, mission type, location, and sortie duration. This study differs from the others in that it focuses on a much finer level of detail, forecasting the demand rate during a particular day. Sherbrooke (1997) found that the sortie number during the day had the highest predictive value of its regressors.[12] The author found that the demand for spares from early sorties was one-third that of the last sortie of the day. The demand for spares from the last sortie of the day was higher due to maintenance groundings and maintenance deferrals from earlier sorties. The author also found that sortie duration did not have a linear relationship with spares demand. In other words, sorties that were twice as long produced less than double spares demand, reinforcing Slay's findings. From Sherbrooke, we test the inclusion of average sortie duration as a covariate in Chapter Four.

Wallace, Houser, and Lee (2000) Found Problems with Proportional Models

Despite improving the ability to forecast costs by using more refined model specification that included sorties and average sortie duration, the Air Force still encountered problems with predicting wartime resource consumption. The Cost per Flying Hour (CPFH) computation predicted significantly larger costs than were actually observed during contingency operations in Operation Desert Storm. Because of the

[11] The Persian Gulf War included two main operations. Operation Desert Shield (7 Aug 1990 - 16 Jan 1991) was a defensive mission to prevent the invasion of Saudi Arabia and move a large number of troops and equipment into theatre. Operation Desert Storm (17 Jan 1991 - 28 Feb 1991) was the operational military campaign to extract Iraqi forces from Kuwait.

[12] During operation Desert Storm, the Air Force flew mission around the clock. The AF counted sorties for a given aircraft, beginning at 0000 hours and ending at 2400 hours.

inaccuracy of the status quo prediction method, the Air Force wanted to develop more realistic forecasting methodology to better allocate budget.

Of significance to our research is the observation that the proportional model appears to work well in certain situations. When there is little variation in both flying hours and associated maintenance needs—as would be the case in non-contingency operations—the parameters are correlated and flying hours appear to be a useful to predicting costs. However, during wartime operations, with associated surge in flying hours, maintenance needs do not increase proportionally. In the case of C-5B, Wallace, Houser, and Lee found that removals (a proxy for cost) only coincidentally correlated with flying hours. In the case of the C-5B, the proportional model predicted maintenance needs, in terms of removals, three times as large as actually occurred. In short, the proportional CPFH model does not predict well with changes—particularly large ones—in the flying hour program.

Wallace, Houser, and Lee (2000) noted that there are factors that contribute to aircraft maintenance in addition to flying hours. In addition to flying hours, they find that the critical parameters to forecasting maintenance needs are ground days, cold cycles (engine start and shut down—a sortie) , and warm cycles (pairs of landings and take off during a sortie where the engines are not shut down). As long as there are small changes in the flying hour program, the proportional model performs well. However, their model outperforms the proportional CPFH model during contingency surges, where flying hours increase dramatically, but landings and maintenance needs do not.

While we maintain that flying hours can be a useful predictor of O&S costs, Wallace, Houser, and Lee (2000) indicate a substantial limitation to the proportional CPFH model. When there are changes in the flying hour program or fixed costs, the proportional model can be an inaccurate method of forecasting. We build on Wallace, Houser, and Lee's critique of the proportional CPFH model by providing an analysis of the implications of using a proportional model in the presence of nontrivial fixed costs in Chapter Six.

Laubacher (2004) Found That Holt's Linear Method Provides Improved AF Budget Estimates

While earlier CPFH-related research identified a variety of different estimating specifications and criticism of the current proportional CPFF model, the Air Force Institute of Technology (AFIT) sponsored several studies designed to increase the accuracy of estimating cost per flying hour (CPFH) for specific MDSs. In building primarily on Wallace, Houser, and Lee (2000), the focus of the AFIT research has been to increase the predictability of MDS-specific CPFH factors through more innovative model specifications. Laubacher (2004) performed the first of the most recent research on CPFH factor estimation improvement.

Laubacher examined three separate forecasting techniques for the Air Force's MH-53J/M, HH-60G, and UH-1N helicopters, with the goal of reducing the differences between forecasted MAJCOM budgets and actual expenses. For the three-year period of FY00-FY02, Laubacher compares actual costs to estimates of CAIG level-one costs (element 1.0 through 7.0 described above) created using a 3-year moving average, the single exponential smoothing method, and Holt's linear method. For five of the eight MAJCOM budgets evaluated, Laubacher's forecasting techniques yielded closer approximations of the actual costs than the budget forecasted by the Air Force. Our estimation approach examines both level-one and level-two CAIG costs, but replaces the above forecasting strategy with econometric models that can accommodate additional parameters, such as flying hours and aircraft age, to the CAIG element budgets.

Hawkes (2005) Found That Last Year's CPFH Rate Is a Good Predictor for This Year's Rate

Hawkes continued the examination of cost per flying hours, but changed the focus of the study from helicopters to airplanes. Hawkes' thesis examined Active and Air National Guard F-16 flying hour data to identify explanatory variables that influence CPFH. Hawkes' research builds linear regression models that predict F-16 CPFH using operational and programmatic variables; he concludes that the previous year's CPFH rate, utilization rate, base location, average age of aircraft, and other calculated variables are statistically significant in predicting CPFH. Notably, Hawkes' primary explanatory variable for this year's CPFH rate was last year's CPFH rate. Hawkes concludes that there is not sufficient empirical evidence to show that average sortie duration (flying hours/sorties) has an effect on CPFH, a finding consistent with his literature review. It is possible that nonlinear sortie duration may have been significant, a hypothesis we will test in Chapter Four.

Armstrong (2006) Found That Marginal CPFH Factors Perform Better Than Proportional CPFH

Armstrong (2006) built on Hawkes' CPFH research efforts by using econometric techniques to build a marginal cost per flying hour model for the F-15. Armstrong used monthly panel data[13] to analyze the

[13] Panel data is the combination of cross-sectional data and time-series data—essentially a cross-sectional analysis over time.

effect of independent variables—including aggregate economic, programmatic, operational, and climatology data—on CPFH rates. Armstrong found significant seasonal trends in the CPFH rates, along with statistically significant variables of average sortie duration and mean monthly temperature difference.[14] Armstrong's model validation indicated that his marginal CPFH model outperforms, in terms of predictive ability, the current proportional CPFH models used by US Air Force.

The Air Force uses a combination of approaches to arrive final budget estimates. Much of the analysis devoted to cost per flying hour research focuses on top-down estimating strategies for specific aircraft platforms. One of the inherent problems with these system-specific methodologies is that they are severely constrained by a lack of cost observations. We will complement the system-specific approaches by including observations from multiple systems in our models, attempting to overcome some of the limitations of previous analyses.

While the above studies focus on the effect of usage on costs, a great deal of research has been accomplished on a similar and related topic, aircraft aging. The following section provides an overview of selected aging aircraft studies.

---

[14] Corrosion research (Guo, 2004) indicated that higher temperatures have a significant corrosive effect on aluminum alloys.

Aging Aircraft

While this study focuses on the relationship between usage and costs, researchers have conducted a substantial amount of research on the relationship between age and costs. The aging of the Air Force inventory has garnered much attention from the Air Force, since most observers react viscerally to the fact that the average age of some of the inventory is twice that of new pilots. Substantial acquisition costs for new systems coupled with budget constraints have forced the Air Force to maintain its inventory far beyond the original design life.

Table 2-1 USAF Aircraft Ages by MD (FY06)

Type	MD	TAI	Average Age
Bomber	B001	68	19.1
	B002	21	12.2
	B052	94	44.8
Fighter	A010	356	25.3
	F015	715	21.4
	F016	1339	17.1
	F117	54	20.7
Recce	E003	33	26.9
	E008	17	6
	RC135	21	43
	U002	27	22.8
Tanker	KC010	59	21.7
	KC135	529	45.5

Type	MD	TAI	Average Age
Special	AC130	22	23.1
	HC130	29	33.7
	MC130	61	32.6
	WC130	10	40.8
Trainer	T001	179	11.9
	T006	243	2.3
	T037	230	41.9
	T038	515	39.5
Transport	C005	112	27.9
	C017	149	5.8
	C021	76	21.7
	C130	483	25.7

Table 2-1 shows the ages and total aircraft inventory (TAI) of a subset of the MDs included in this analysis. As of FY06, the two oldest MDs in the Air Force inventory—the B-52 and KC-135—were both well over 40 years old, on average.[15] The presence of extremely aged platforms, along with the documented maintenance demand "bathtub curve"[16] that states that platforms cost more to maintain as they age, supports the need to control for age when examining the relationship between aircraft usage and maintenance costs.

[15] Reliability & Maintainability Management Information System (REMIS)
[16] Foster (1983) describes the systems bathtub curve as having higher failure rates during the initial burn-in period and at the end of life as many parts fail due to wear.

Pyles (2003) Found Significant Age Effects on Maintenance Requirements

Pyles (2003) authored a comprehensive study of age effects on workload and material consumption of Air Force aircraft. He focused on distinguishing aging effect from a variety of comingled events, such as changes in the accounting systems, changes in maintenance system personnel and organizations, and changes in maintenance procedures. Among many age-related research questions, Pyles specifically addressed, by platform, how rapidly costs grow with age and how those cost growth rates might change in the future.

Pyles found that maintenance requirements increase as aircraft age. Although growth differed across fleets, flying hours, and flyaway costs, he generalized that the more expensive, more complex aircraft experience higher growth rates; Pyles noted that the key factor affecting the growth rate was flyaway costs, with growth being proportional to said cost. Additionally, Programmed Depot Maintenance (PDM) workload growth accelerated after thirty years of service and the workload grew more rapidly in more expensive aircraft.

Keating and Dixon (2003), and Keating, et al (2005) Determined Optimal Replacement Dates for Aging Aircraft

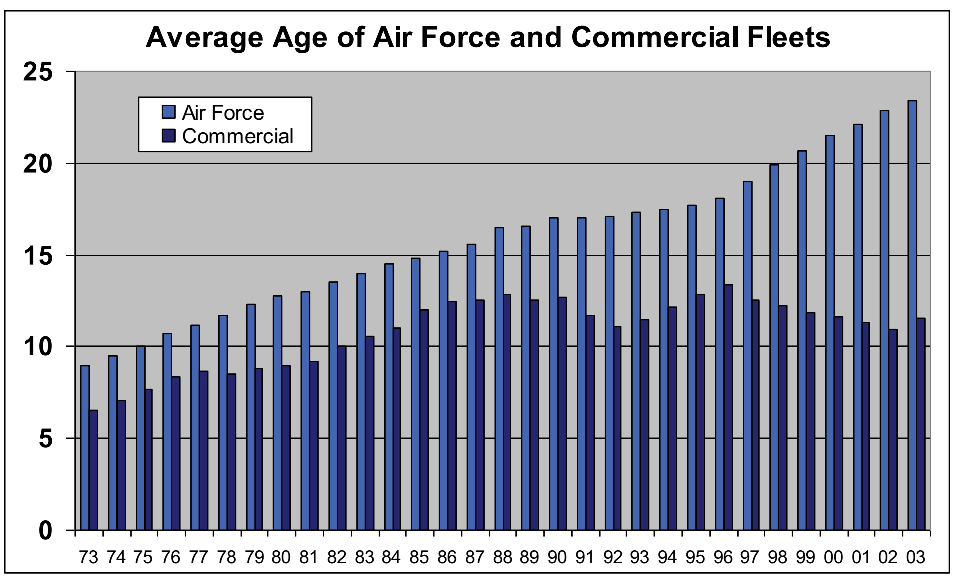
Building on the Pyles' aircraft aging work, Keating and Dixon (2003) focused on the problem of determining when to replace aging aircraft. Their research introduced an approach for deciding when it is optimal to replace an aging system; the Air Force should retain a particular fleet until the incremental costs exceed the average cost per available year of a replacement system. In their initial effort, they focused on data from the C-21A transport aircraft and the KC-135 tanker aircraft. They were able to use their model to determine optimal replacement dates for both fleets. For example, the Air Force should retire the C-21A and the KC-135 in about 2020 and about 2010, respectively.

In a 2005 RAND report, Keating, Snyder, Dixon and Loredo extended the findings of Keating and Dixon (2003), showing that their model can be employed to evaluate modifying an aircraft to extend its life, instead of retiring it. Based on cost criteria, their approach indicates whether the Air Force should extend the life of a particular platform through modification, or whether an aircraft should be retired before the modification. They specifically applied the model to the C-5A cargo aircraft, to evaluate the cost implications of a planned modification. They extend their model by including depot-level capacity as a choice variable. They find that increasing the C-5A depot capacity could reduce maintenance costs, as forcing a highly-valued system to wait for maintenance is extremely expensive. Their research uses age-related, increasing maintenance costs as a key element in the analysis of repair and replace decisions. Older platforms are associated with higher annual maintenance costs, which eventually rise to a level that favors replacement over additional repair.

Dixon (2006) Found That Aircraft Maintenance Costs Increase as Commercial Fleets Age

A recent examination of commercial aircraft showed that the total maintenance cost increases at a decreasing rate as commercial fleets age; in other words, the age effect decreases as the fleets get older, approaching near zero growth. While the rate at which maintenance costs grow slows as platforms age, total maintenance costs may continue to increase. Dixon showed a commercial aircraft age effect (log(total costs/flying hours)) ranging from over 15% for a one-year age increase for aircraft under six years old to less than 1% for aircraft over 12 years old. This study also showed that for commercial platforms, the type of aircraft and the carrier (owner) do not have a statistically significant effect on age effects.

Figure 2-2 Average Age of USAF and Commercial Fleets



However, Figure 2-2 shows that there may be problems with generalizing aging effect results from commercial data to Air Force platforms. From 1973, not only have Air Force platforms been older on average than the commercial fleets, but since about 1982 the Air Force inventory has been older than the commercial fleet has ever been. Certain types of Air Force fleets, such the B-52 and KC-135, are well beyond the range of commercial data. As of 2003, the Air Force inventory at an average of 23.4 years was more than twice as old as their commercial counterparts at 11.6 years. As time progresses, the average age of the Air Force inventory continues to grow well beyond the range of commercial data, possibly limiting the usefulness of comparisons. Moreover, differences in flying hour variability—commercial fleets fly very consistently, while military flying hours can be quite variable—and total annual flying hours (civilian platforms fly many more hours per tail annually) may limit the applicability of commercial data to military analyses. Further, commercial carriers intentionally avoid large increases in maintenance costs seen for military aircraft by removing aging tails from their inventory.

Aircraft age is not the only driver of O&S costs, and may not be the primary driver. However, the body of literature on aircraft aging suggests that aircraft age can account for a nontrivial portion of the increase in maintenance costs; it is reasonable to control for age when investigating how usage relates to O&S costs. We control for age in all of our models, finding that the age parameter typically has a significant impact on the fitted cost values.

From the above literature review, two themes emerge. First, and most importantly, there are documented limitations to the proportional CPFH metric currently used by the Air Force. Second, there is some disagreement on whether age and average sortie duration are statistically significant when modeling the effect of flying hours on O&S costs. In order to examine both of these themes in greater detail, we discuss the data and systems associated with costs and usage in the next chapter.

CHAPTER THREE

Data Overview

This chapter discusses the data collected to analyze the relationship between flying hours and usage. We group the data into three data types: identification, cost, and usage. Two Air Force data systems provided the data necessary for the analysis. Cost data, broken out by the CAIG cost element structure, were collected from the Air Force Total Ownership Cost Decision Support System (AFTOC) data system. Usage data, including flying hours, landings, and aircraft age data were collected from the Reliability & Maintainability Management Information System (REMIS) system. We related the cost data to the usage data with common identification fields such as Mission Design and Mission Design Series. The following sections discuss the fundamental properties of the data collected and the process of aggregating the disparate datasets. This discussion will provide insight into the relationship between the variables and begin to address the strategy of the modeling effort.

Identification Variables

The first set of variables includes those that describe the individual observations— termed the identification variables. Included in this set are the observation number, fiscal year (FY), aircraft type (type), aircraft mission design (MD), and aircraft mission design series (MDS). Aircraft types include broad categories of aircraft such as bomber, fighters, and tankers. Aircraft mission design provides a more specific description of the aircraft (F-15, F-16, C-5). One level below MD, aircraft mission design series provides the most specific description of aircraft supported in the majority of Air Force data systems (F-15A, F-15B, F-15E). There are further levels of military aircraft type nomenclature, such as aircraft block, but this level is not supported by most comprehensive AF cost reporting systems. Our analysis will focus on the MD level, since cost data are not collected by tail.

REMIS Usage Data

We collected all of the usage data from REMIS, a system of systems that provides information on a wide variety of aircraft maintenance aspects. As described by the AF Portal, "REMIS provides authoritative information on weapon system availability, reliability and maintainability, capability, utilization, and configuration." [17] REMIS provides, by tail, sorties, flying hours, takeoffs, landings, aircraft production

[17] AF Portal: https://www.my.af.mil/gcss-af/afp40/USAF/ep/index.do?command=application#R

date, and other maintenance data. The source of the MD and MDS information is the REMIS System, but aircraft Types were assigned from AFTOC, separating the information on Tankers and Transports.

REMIS provides monthly data on several variables that suit the purposes of this analysis: flying hours, sorties, possessed hours, and landings. Since cost data reporting systems maintain historical costs only for complete fiscal years, however, this study limits its resolution of usage data to fiscal years. Flying hours and landings are self-explanatory. Sorties pertain to the number of missions flown by an aircraft—a sortie may include multiple takeoffs and landings. The Possessed Hours variable is the number of hours an aircraft is "owned" by a particular installation. Possessed hours for a single aircraft should total to about 8760 for a given fiscal year (365 x 24) across all of the owning installations. The total possessed hours for an aircraft in a leap year is 8784. There are situations in which an aircraft may have less than 8760 hours per fiscal year, the typical reason being retirement.

The Possessed Hours variable in itself may not obviously relate to maintenance costs, but it does prove useful in calculating an approximation for total aircraft inventory. Removing the possessed hours for Davis Monthan[18] leaves an approximate number of active aircraft hours. Dividing the number of active hours by 8760 (or 8784 in a leap year) gives an approximation of the number of tails active during a given fiscal year—typically referred to as Total Active Inventory (TAI). This approximation tracks quite well with the published numbers for TAI, including those in AFTOC, and is sufficient for this analysis. Flying hours per TAI (FH/TAI), flying hours per sortie (FH/sortie), and flying hours per landing (FH/landing) can be calculated from the above information; these fields give measures of the intensity of activity for a given MDS or MD.

In the REMIS there are multiple instances of tails that do not have flying hours, but do have possessed hours for a given fiscal year. One thousand twenty five tails have at least one occurrence of zero annual flying hours in the FY96-FY06 period. The majority of these instances occur in F-16, T-37, and T-38 MDs. There are a total of 15.4 million possessed hours that do not have associated flying hours, less than 2.5% of the total of 617.1 million possessed hours. There are 6795 tails in this unrestricted dataset that includes tails that we considered for removal. Ultimately we decided to remove only 304 tails that have

[18] Located in Tucson, Arizona, Davis Monthan AFB—also known as the "boneyard"—is home to the Aerospace Maintenance and Recovery Center (AMARC). The Air Force and other services store aircraft at AMARC, due to the advantageous climate. Aircraft in storage may still report possessed hours in REMIS, but these hours are excluded from our dataset since they do not have associated flying hours or O&S costs.

significant possessed hours and less than 120 flying hours during the entire 11 year period. There are 255 tails with less than 50 total flying hours during the same period and 139 tails with no flying hours.

The removal of these tails from our dataset had a negligible impact on the flying hour specifications. However, the accrual of possessed hours against no flying hours does affect the calculation of TAI and related calculations. Ultimately, the concern is to be able to properly test for whether costs vary with flying hour or with TAI, so it is important to calculate TAI properly and consistently.

The difficulty in removing these tails is that the Air Force still shows them on the books in REMIS. These aircraft are not officially terminated and have not been relegated the boneyard at Davis Monthan. Additionally, it is possible that these aircraft still contribute to O&S costs, even with a few or no flying hours. Unfortunately, given that the cost data are reported at the MDS level, it is impossible to tell if a given tail contributes to O&S costs. Table 3-1 provides an example of the type of data available in REMIS.

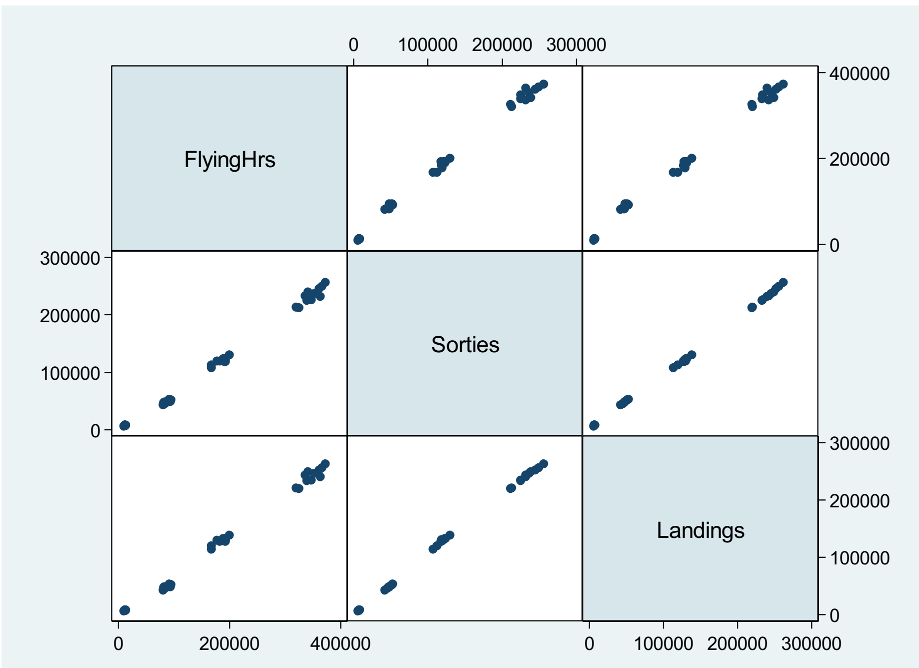
Table 3-1 : Example Raw Data from REMIS

MDS	Tail Number	FY	Installation	Flying Hours	Sorties	Possessed Hours	Landings
F016A	8000000547	2005	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	2006	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	2002	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	2003	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	2004	AZRAQ HIGHWAY STR	0	0	8784	0
F016A	8000000547	1998	AZRAQ HIGHWAY STR	0	0	7016	0
F016A	8000000547	1998	HILL AFB UT	4.7	5	1744	8
F016A	8000000547	1999	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	2000	AZRAQ HIGHWAY STR	0	0	8784	0
F016A	8000000547	2001	AZRAQ HIGHWAY STR	0	0	8760	0
F016A	8000000547	1997	HILL AFB UT	0	0	8743	0

The minimum of 120 flying hours seems reasonable, since it eliminates the aircraft with both the extremely low flying hours and also those with very low FH/year (10FH/FY or less). This action removes 304 tails, 5468.5 flying hours, and 7,284,231 possessed hours. The final dataset includes information from 63,995 tail-level observations, 22,112,505 flying hours, 546,907,577 possessed hours, 34MDs and 61 MDSs.

Age is another calculated field. REMIS provides tail-level acceptance date and termination date, both necessary to calculate tail and average ages. A given tail will age one year every fiscal year. However, the average age for the tails within an MDS or MD varies, depending on terminations and accessions. Therefore, the average age for MD and MDS must be calculated from the tails extant in that fiscal year, accommodating tails active for less than a complete fiscal year. For example, if a tail enters the AF inventory halfway through the fiscal year, its value to the average age calculation is 0.5 years. The calculated average ages of the included MDSs tracks well with the published information on aircraft age. We also calculated the logs of each of the variables, except age. A model specification that includes log(age) would result in an uncommon interpretation, since changes in age are not generally referred to as percentages. For example, it is clearer more typical to reference an age increases of 5 years rather than an age increase of 10%.[19]

Figure 3-1 Correlation between FHs, Landings, and Sorties for Fighters

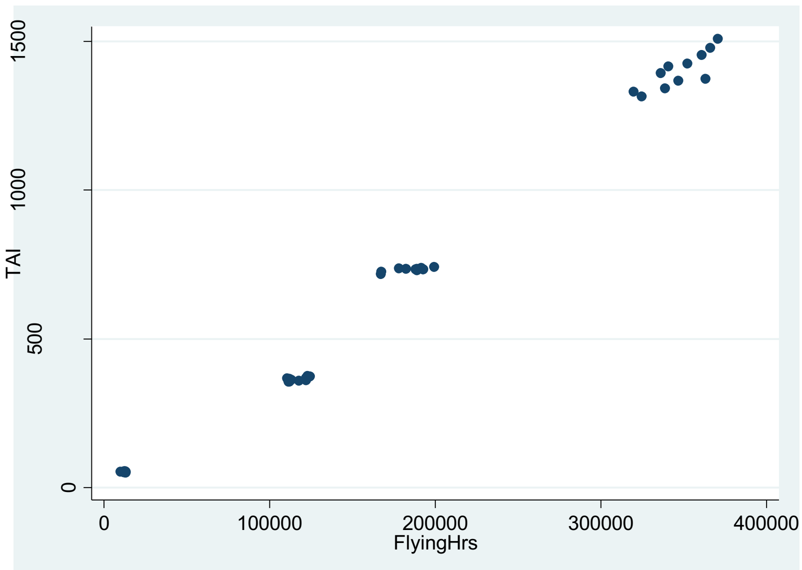


Most of the usage data have the potential for strong correlation. This correlation is especially evident when examined by MD or type. Because of the strong correlation between these variables, they are likely to cause multicollinearity if used simultaneously in a model. TAI is calculated from possessed hours, so their strong correlation is unremarkable. Figure 3-1 shows the relationship between flying hours, sorties and landings for fighters, but a similar pattern exists for the other types. Because of this

[19] Other studies that employ specifications with log dependent variables and linear age include Hildebrandt (1990), Dixon (2006)

strong correlation, sorties and sortie dependent calculated fields will be excluded from flying hour model specifications.

Figure 3-2 Correlation between FHs and TAI for Fighters



Flying hours and TAI are the two primary variables of interest, since most cost forecasting is performed using one or both. However, due to their very strong correlation depicted for fighters in Figure 3-2, it may not be advisable to have a specification that includes both simultaneously. In Chapter Five, therefore, we will use two model specifications—one with FHs and the other with TAI—to test the ex ante Pentagon cost categories of cost per FH, cost per TAI, and Fixed Costs.

AFTOC Cost Data

AFTOC is the Air Force's primary weapons system cost reporting database. AFTOC, which focuses primarily on O&S data, integrates data from a variety of accounting, budget, and supply systems. For the purpose of this analysis, AFTOC provides normalized, annual cost of ownership data by MD. AFTOC can provide data at the MDS-level, but most of AFTOC's costs are allocated (not collected directly in the category represented), limiting the usefulness of analysis below MD-level.

The Air Force collects and reports financial data in a wide variety of formats. Due to congressional requirements for budgeting, the Air Force distinguishes between budget dollars, expended dollars, and a variety of other types of money. Our analysis of O&S costs is performed at an aggregated level, rather than by tail, so we will discuss O&S costs as a whole. However, it should be noted that, O&S costs comprise a number of different appropriations and that each face different congressional budget rules.

It is useful to separate the O&S data, not by appropriation, but by AF Cost Analysis Improvement Group (CAIG) cost element structure. The CAIG costs structure is the breakout used as the basis for cost analysis, cost estimating, and budget work within the Air Force and other services. The CAIG elements consist of seven level-one categories, as discussed in Chapter One. Each of these level-one categories contains at least one level-two subcategories, the level at which we will conduct the analysis, where sufficient data permits. [20]

However, AFTOC does not collect budget data by CAIG category, but rather by Program Element Code (PEC)—a budget category that groups cost for a certain weapon system or systems. The Air Force collects budget data via PEC, and then allocates dollars to MD and MDS. Budget information for MD and MDS, therefore, are approximations of costs for those areas. The AFTOC home page provides a master cross-reference table that shows the specifics of the allocation from PEC to MD (Lively, 2007). We discuss the implications of this allocation below.

Additionally, the Secretary of the Air Force for Financial Management (SAF/FM) organization groups the level-two CAIG costs into three categories for budget estimating purposes: costs that vary with flying

[20] A thorough explanation exists in Appendix B of the CAIG O&S Support Guide http://www.dtic.mil/pae/

hours, costs that vary with Total Aircraft inventory (TAI), and fixed costs. We analyze the validity of these ex ante groupings and suggest possible improvements in Chapter Five.

Dataset Structure, Abridged

Table 3-2 shows a subset of the data used to construct the econometric models. In general, there are 11 years of usage and cost data at the MD level. Some MDs, including the C-141, E-8, and T-6, have less than 11 years of data—typically if they are newer or older platforms. Certain observations may have been omitted if there was no associated cost data, but that was rare. The final MD level dataset includes 361 observations for 34 MDs.

Table 3-2 Example Dataset for the B-1

Identification					REMIS Usage				AFTOC Cost
PID	**FY**	**Type**	**MD**	**MDS**	**FlyingHrs**	**Age**	**TAI**	**FH/TAI**	**CAIG Tot**
1	1996	Bomber	B001	B001B	26452.1	9.4	95.3	277.7	1040362089
1	1997	Bomber	B001	B001B	24750.7	10.4	95.0	260.6	961016643
1	1998	Bomber	B001	B001B	23737.4	11.4	93.4	254.1	1065103121
1	1999	Bomber	B001	B001B	22883.1	12.4	93.0	246.1	1018833974
1	2000	Bomber	B001	B001B	24646.4	13.4	93.3	264.3	1100394982
1	2001	Bomber	B001	B001B	24570.8	14.4	93.0	264.2	1138445795
1	2002	Bomber	B001	B001B	25970.5	15.4	90.8	285.9	1261553132
1	2003	Bomber	B001	B001B	20832.9	16.2	71.0	293.4	1135463989
1	2004	Bomber	B001	B001B	27463.7	17.1	67.3	408.3	1201495845
1	2005	Bomber	B001	B001B	21208.8	18.1	68.0	311.9	1067092127

Data Aggregation and Allocation

There are two separate, but important issues surrounding the aggregation of data. The first deals with the ability to aggregate tail level usage data to higher levels; the second issue deals with the disaggregation of MD-level cost data to lower levels. Both of these problems could potentially impact the analysis and require some discussion.

Most of the AFTOC and REMIS fields are additive in the sense that one can accumulate them at a higher level of aggregation by simply adding them together. However, variables that contain average values, such as age, cannot be averaged at the higher level, since it may distort that actual value. Constructed values, such as flying hours per sortie are also not additive. Each of the calculated fields in the dataset must be constructed properly at the chosen level of aggregation—for this study MD-level. For example, the age variable is the average of all of the tail ages at the MD-level and flying hours per sortie averages the individual tail flying hours per sortie value at the MD-level.

The second data aggregation issue, briefly discussed in Chapter One, is a data collection limitation rather than a calculation issue. We explained above that AFTOC, in some cases, allocates MD budget data to MDSs using flying hours. There are two separate allocation issues: the allocation of costs from Program Element Code (PEC) to MD and allocation of MD costs to MDS. Since the Air Force collects costs in PECs which may comprise more than one MD, the PEC costs must be allocated to the appropriate MD; costs may not be properly assigned to a given MD. It follows that if there are allocation problems at the MD level, they will be passed to the MDS level. Secondly, some costs are allocated from the MD level to the MDS level by proportion of flying hours. An MDS specification would likely overstate the relationship between O&S costs and flying hours.

Tables 3-3 and 3-4 provide an example of this allocation for CAIG element 1.2 maintenance personnel costs for the C-5 and F-16. While narrow in scope, this example underscores the need to conduct our analysis at the MD level rather than the MDS level. It is particularly striking, since CAIG 1.2 costs are associated with total aircraft inventory rather than flying hours; we expect CAIG 1.2 costs to correlate with TAI, but in this example they correlate with flying hours.

Table 3-3 shows the MDS FY06 raw values for CAIG 1.2 costs, flying hours, and TAI for the C-5 and F-16. Immediately below, Table 3-4 shows the ratio of the MDS values to the total MD values for each of the variables. For example, the C-5A flew 22,790 hours of the total 52,982 C-5 hours for a ratio of 22790/52982 = 0.4301.

Table 3-3 MDS Allocations for FY06

MD	MDS	CAIG 1.2 Costs	FHs	TAI
C-5	C-5A	152090357	22790	65.5
C-5	C-5B	195971786	29759	47.8
C-5	C-5C	4923638	433	1.8
F-16	F-16A	21686550	5800	166.5
F-16	F-16B	24963977	5863	56.0
F-16	F-16C	1206069618	271586	1071.8
F-16	F-16D	189322285	37871	182.8

Table 3-4 Ratios of MDS allocations to Totals for FY06

MD	MDS	CAIG 1.2 Costs	FHs	TAI
C-5	C-5A	0.4309	0.4301	0.5696
C-5	C-5B	0.5552	0.5617	0.4152
C-5	C-5C	0.0139	0.0082	0.0152
F-16	F-16A	0.0150	0.0181	0.1127
F-16	F-16B	0.0173	0.0183	0.0379
F-16	F-16C	0.8364	0.8457	0.7256
F-16	F-16D	0.1313	0.1179	0.1237

While the cost of maintenance personnel is typically assumed to be correlated with the number of aircraft—TAI—and not flying hours, it can be seen by comparing the CAIG 1.2 cost column and the FHs column in the above tables that there is a strong association between the two quantities at the MDS level. CAIG 1.2 costs appear to be allocated from the MD level to the MDS level by flying hours for both the C-5 and F-16 in FY06; the ratios in the CAIG 1.2 cost column more closely resemble the ratios in the flying hour column than the TAI column. For example, for the C-5A cost ratio of 0.4309 is substantially closer to the flying hour ratio of 0.4301 than the TAI ratio of 0.5696. A single similarity between a cost ratio and a flying hour ratio could be the result of coincidence, but the similarities occur across the

MDSs. Moreover, the lack of correlation between the cost and TAI ratios provides further evidence that costs were allocated by flying hours. We deduce from the above relationship and discussions with analysts familiar with AFTOC that a nontrivial amount of costs within AFTOC are allocated from the MD level to the MDS level by flying hours.

In Chapter Five, we show that CAIG 1.2 costs increase by approximately 56% as flying hours double using the MD-level specification. Had we modeled these costs at the MDS-level, our results would have been biased towards the proportional model—the 0.56 coefficient on flying hours closer to one. Since we are uncertain that O&S costs were properly allocated to the MDS level, we choose to use the MD aggregated data for the baseline models.

Implications of Data Selection to Modeling Goals

The primary goal of our analysis is to establish a top-level relationship between O&S costs and flying hours. A refinement of this goal is to establish insight into the Cost per Flying Hour (CPFH) metrics used to generate Operations and Support (O&S) budgets and to comment on the current SAF/FM structure of O&S costs. The policy goal of this research is to create efficiencies in the Air Force budgeting process by enabling planners to better allocate scarce resources.

In terms of the data themselves, there are several overarching issues that need to be addressed. First, due to the high degree of correlation between many of the variables, variables will be selected to addresses multicollinearity and eliminate redundancy. For example, aircraft usage data such as flying hours, sorties, and landing exhibit strong correlations; a regression model including all of them will display a high degree of multicollinearity.

Secondly, modeling the relationship between cost and usage involves tradeoffs between different levels of data aggregation. The dependent variable of cost is available only on an annual basis dating back to fiscal year 1996 in AFTOC—allowing for 11 observations per MD or MDS. AFTOC reports costs at either the MDS or MD level, but there appear to be significant cost allocation problems at the finer MDS allocation. The trade between MD and MDS is a lower degree of allocation problems at the MD level versus greater statistical power and more observations at the MDS level. However, given that the allocation of costs to the MDS level is likely biased towards the proportional model, we should be conservative in our approach and create MD-level models.

In addition to the MD-MDS trade, the cost data can be decomposed into Cost Analysis Improvement Group (CAIG) cost elements. The CAIG elements fall into seven top-level categories (level-one) and 22 sub-elements (level-two). Lower levels of cost aggregation allow the model to more specifically describe the relationship between a particular cost and usage, but may cause statistical issues due to lack of data in some categories.

On the usage side, the data are collected by an individual aircraft or tail. Since AFTOC reports costs at the MD or MDS level, usage data need to be appropriately aggregated to align with the cost data. It would be preferable to be able to create models at the tail level to increase statistical power. This specification is possible with certain dependent variables, but not with cost. The aggregation of the cost variable forces the aggregation of usage data to the MDS level, at a minimum. As a result of possible cost allocation issues, we selected MD-level model as baseline specification level of estimation.

Information derived from the models must be able to address the mechanics of CPFH calculation—marginal versus average costs—and provide sufficient information to evaluate the SAF/FM CAIG cost grouping of cost per FH, cost per TAI, and Fixed costs. Overall, the paucity and structure of O&S cost data forces uneasy modeling decisions. To address the research question in the following three chapters, we chose the log-log specification at the MD-level as the best modeling approach, given the large variance in MD-specific data and substantial limitations of the cost data. In Chapter Four, we discuss the log-log specification in further detail in our examination of the relationship between flying hours and O&S costs.

CHAPTER FOUR

USAF Aircraft Usage and O&S Costs

In general, military decision makers would benefit from improving their understanding of the relationship between usage and maintenance costs. Derived from that relationship, improved estimates of CPFH factors would enhance the ability of the Air Force to estimate future budgets. In other words, improving our understanding of the usage effect will enhance estimates of future maintenance costs, thereby informing future O&S policy decisions. In this chapter we will discuss the top-level relationship between O&S costs and flying hours. In order to inform improvements the CPFH factors in Chapter Six, we begin by investigating the usage effect in three separate regression model specifications: indicator, average sortie duration, and flying hour per total aircraft inventory. Also, we will briefly discuss a performance specification, which replaces the aircraft-type indicator variables with variables such as weight and thrust.

Econometric Methodology Overview

The bulk of the methodology involves analysis of relationships between a variety of independent variables and O&S costs. The central portion of the empirical analysis is to separate the costs associated with aircraft usage from calendar costs of aging, while controlling for platform-specific variation. The principle benefit of this analysis is investigating the functional form of the Cost per Flying Hour (CPFH) metric, which directly relates to O&S budget formulation. Chapter Five will refine and extend the baseline modeling by identifying the relationship between usage and disaggregated costs—level-two CAIG elements—to identify to what degree the elements are variable or fixed costs.

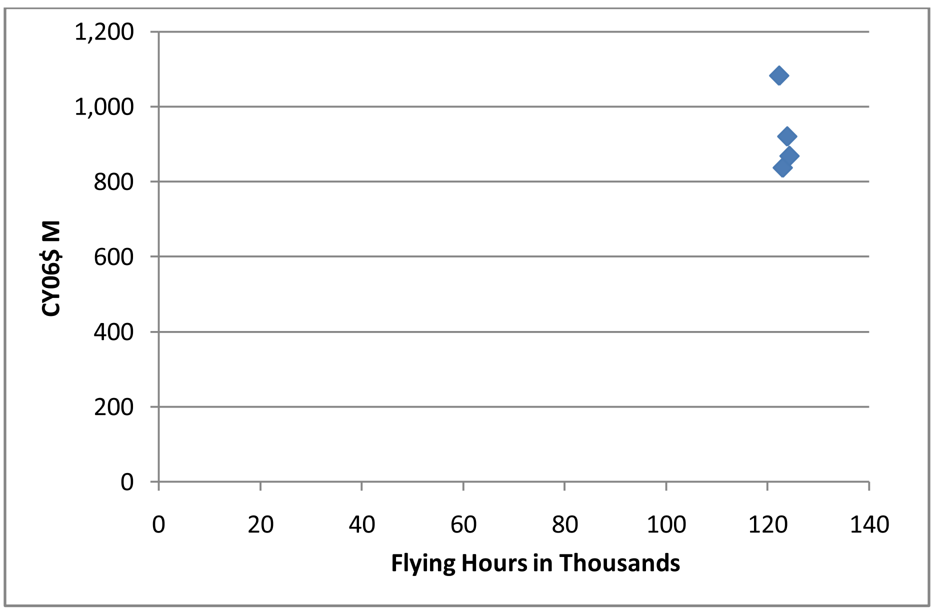
Initially, we will combine cost data from the AFTOC database with the REMIS usage data.[21] As discussed earlier, we intend to use cost data at the MD level, so the usage data will need to be aggregated to ensure proper allocation. While this aggregation limits the statistical power of this analysis, the addition of MD-level maintenance expenditure data allows us to analyze the extent to which inflation-adjusted costs (in CY06$) changed with usage.

[21] While REMIS data are available to at least FY94, AFTOC data are limited to the FY96 to FY06 range.

Variation in Flying Hours

In order to address the effect of usage on O&S funding, we need to ensure that there is sufficient variation in the aggregate flying hours. Historically, and particularly in peacetime operations, there has not been enough variation in flying hours to adequately identify a usage effect; the advent of contingency operations in Iraq and Afghanistan provides a natural experiment with which to investigate usage effect. Figure 4-1 illustrates the difficulty in estimating a usage effect without variability in flying hours—little or no variation may provide for no significant relationship between flying hours and costs. The variation in costs in the below figure is attributable to something other than just flying hours.

Figure 4-1 Flying Hours from FY96 to FY99 for the A-10



Inflation and AFTOC

In order to compare costs across fiscal years, one must mitigate the effect of inflation; budget dollars in FY96 are not equivalent to budget dollars in FY06. The Air Force terms budget dollars—the funding required to be spent in a given fiscal year—"then year" dollars; this is equivalent to the nominal dollars concept. When determining future budgets it is appropriate to discuss funding in terms of the actual dollars that will be needed—TY$ are suitable. However, when comparing funding between fiscal years, one must control for the effect of inflation. The equivalent of real dollars in Air Force terminology is either constant year (CY) dollars or base year (BY) dollars. The constant year or base year terminology is equivalent and interchangeable. When discussing funding in either BY or CY, a fiscal year (FY) must be included to show the year to which all of the dollars are adjusted. In this study, we us CY06$ for all of the regression models.

However, there are many different ways to adjust for inflation, to convert the financial data to CY$. In this study, we use the inflation rates sanctioned by the Office of the Secretary of Defense (OSD) and incorporated into the AFTOC system. One of the problems with the OSD inflation rates is that they have historically been considered too low—especially in specific areas such as contractor pay. However, the CPI for the analysis period is also fairly low. For comparison, Table 4-1 shows the OSD weighted rates compared to the CPI. There is no appreciable difference between the CPI and OSD inflation rates—with the OSD rates having the advantage of including a weighted index (accounts for multi-year outlays)—so there is no substantive reason to switch between OSD rates and CPI. Therefore, the baseline models will use the AFTOC supplied CY06$ CAIG information.

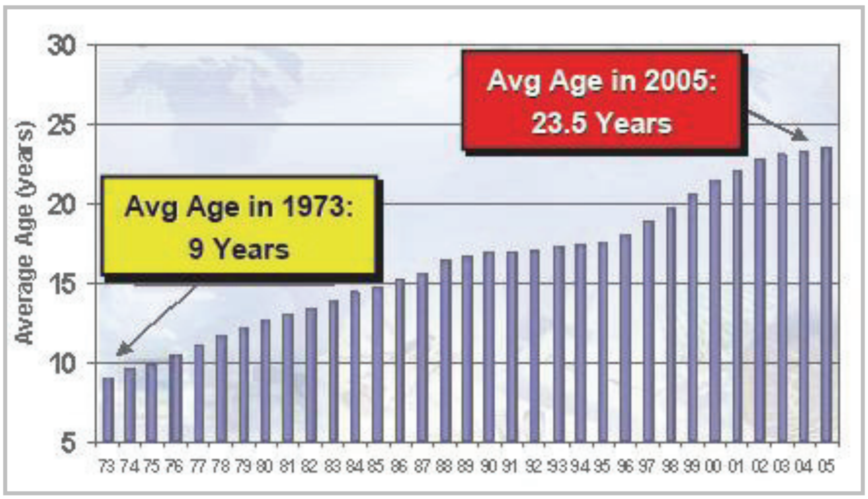
Table 4-1 Comparison of CPI to OSD Inflation Indices (BLS, 2007)

	CPI	**OSD**	**OSD**	**OSD - CPI**	**OSD - CPI**
FY	Annual	MilPay	Aircraft Procurement	MilPay Delta	Aircraft Delta
1996	**1.0300**	1.025	1.020	-0.006	-0.010
1997	**1.0215**	1.029	1.021	0.007	-0.001
1998	**1.0149**	1.029	1.007	0.014	-0.008
1999	**1.0263**	1.034	1.008	0.008	-0.018
2000	**1.0345**	1.045	1.014	0.010	-0.021
2001	**1.0265**	1.040	1.018	0.013	-0.008
2002	**1.0151**	1.061	1.008	0.046	-0.007
2003	**1.0232**	1.053	1.010	0.029	-0.013
2004	**1.0254**	1.043	1.020	0.018	-0.005
2005	**1.0469**	1.037	1.028	-0.010	-0.019
2006	**1.0206**	1.032	1.031	0.011	0.010

Controlling for Aircraft Age

While this research focuses on the relationship between usage and O&S costs, the origin of this type of study lies in the realm of aging aircraft. The United States Air Force began to consider the issue of aging aircraft with a great sense of urgency in the early 1990s. At that time, Air Force leadership recognized that the costs for maintaining aircraft were growing and, simultaneously, the average age of its fleets was steadily rising. Figure 4-2 illustrates the aggregate aircraft aging problem faced by Air Force planners: the average age of USAF aircraft inventory has more than doubled over the last 30 years.[22]

Figure 4-2 Average Age of USAF Aircraft



We include the Age variable to assure the reader that there are no obvious omitted variable bias (OVB) issues with the model specification. Most of the literature refers to an aging effect, but this research focuses on a usage effect while controlling for age. In some specifications, the age variable is not significant. Where the age variable is significant, its coefficient is very small. Regardless of the seemingly small coefficients show in our results, Age appears to have substantial contribution to predicted costs in the more aged platforms.

Much of literature on age shows that it is likely that the age effect varies with the age of the aircraft (Pyles, 2003 and Dixon, 2006). This might warrant the inclusion of an age^2 variable or breaking up age into separate groups—0-10 years, 10-20 years, etc. Because of the potential with overspecification in the indicator models, it might be preferable to include just the age variable, possibly with the Age^2 or

[22] Faykes, Major General Frank, *FY07 Air Force Budget Briefing*, 2005.

other transformed variable, depending on diagnostics. We tested the inclusion of the additional costs variable described above and found, in most cases, that the age variation was nearly completely controlled for by the original age variable.

It is possible that the aggregation of data may mask a portion of the aging effect. Particularly at higher levels of aggregation—such as MD—the use of average age in an observation may mask the actual aging effect by including new aircraft in the average. This is one of many aspects of cost modeling that would be greatly aided by higher fidelity level cost models.

Indicator Specification

We will explore several different model specifications relating O&S costs to flying hours. In selecting a model, we will pay close attention to both the associated statistical diagnostics, as well as the explanatory facility of the model. We can use ordinary least squares (OLS) with MD (for example F-15, F-16) fixed effects to estimate age and usage effects: $Cost_{my} = \alpha + \beta * Age_{my} + \delta * Usage_{my} + \mu_m + \varepsilon_{my}$ where m is the aircraft mission design or MD (e.g. F-15, F-16), y is the year, α is the intercept, β is the age effect, δ is the usage effect, μ_t represents the fleet fixed effects, and ε is the error term. We avoid year fixed effects by using constant year dollars in the cost variable and including the average aircraft age.[23] We initially believed that certain O&S costs may need to be the delayed by one or more years—last year's flying hours affect this year's costs—but discussions with maintainers and lack of significance of the lag variable in the statistical models indicated that aircraft usage would likely affect most costs in the same fiscal year.[24]

We use log-transformed variables to control for wide variation of the cost and usage data within and between the MDs.[25] The log-log specification will perform better with heteroskedastic data since it is a variance controlling transformation; the interpretation as a constant elasticity also aids in the interpretation of a common flying hour coefficient across MDs. The baseline specification will be: $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + \mu_m + \varepsilon_{my}$. The log-log specification has the advantage of

[23] We tested the viability of including fiscal year fixed effects, but found that the age variable already accounted for much of the variation explained by the ten FY indicators.

[24] Armstrong used a lagged cost variable, warranting further investigation of this issue. Interviews with maintainers indicated that wing maintenance is done by flying hours (e.g. every 300 hours) and depot maintenance by calendar.

[25] All of our models use log base e, the natural logarithm.

constant elasticity interpretation and is supported by much of the aging aircraft literature, which uses log-based models.

Baseline Indicator Specification

Total O&S costs (w/o 2.1 fuel); MD Indicators with Age

The below results derive from the unconstrained MD model—all 34 MDs included, omitting the MD indicator for C-130. We omit the C-130 since it has costs within all of the level-two CAIG costs elements and is, therefore, useful as an omitted category for each of the disaggregated cost models. We run the regression with robust standard errors to mitigate problems with heteroskedasticity. The below model shows that the coefficient on Ln(flying Hours) is 0.56; as flying hours double, O&S costs excluding fuel increase by 56 percent.

Indicator Specification: $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + \mu_m + \varepsilon_{my}$.

Table 4-2 Baseline Model Results

```
Linear regression                                      Number of obs =     361

                                                       F( 35,   325) = 1278.66
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9729
                                                       Root MSE      =  .25232

------------------------------------------------------------------------------
             |               Robust
 lncaigtot21 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 lnflyinghrs |   .5567066   .0701463     7.94   0.000     .4187086    .6947046
         age |   .0580409   .0047241    12.29   0.000     .0487472    .0673347
        a010 |  -.0904037   .0609626    -1.48   0.139    -.2103349    .0295275
       ac130 |   -.187737   .2321092    -0.81   0.419    -.6443632    .2688892
        f016 |   1.030994   .0676762    15.23   0.000     .8978556    1.164133
        f117 |   .2649531   .2199092     1.20   0.229    -.1676722    .6975783
       kc135 |   -.945577   .0893918   -10.58   0.000    -1.121437   -.7697174
       _cons |    13.2254   .8848913    14.95   0.000     11.48456    14.96624
```

Table 4-2 shows truncated STATA results for the baseline indicator model; the full model is included in the Appendix. OLS appears to be appropriate since the model does not appear to violate the three basic assumptions necessary for linear regression: ε's are independent between observations, ε's are normally distributed, and Var(ε) is constant.[26] The model F-test shows Prob > F = 0.0000 indicating that

[26] During the modeling process we used standard statistical diagnostic tools such as Q-Q residual plots, plotting residuals versus fitted values, and other techniques to evaluate leverage and influence. The diagnostics were unremarkable in terms of real-world data modeling.

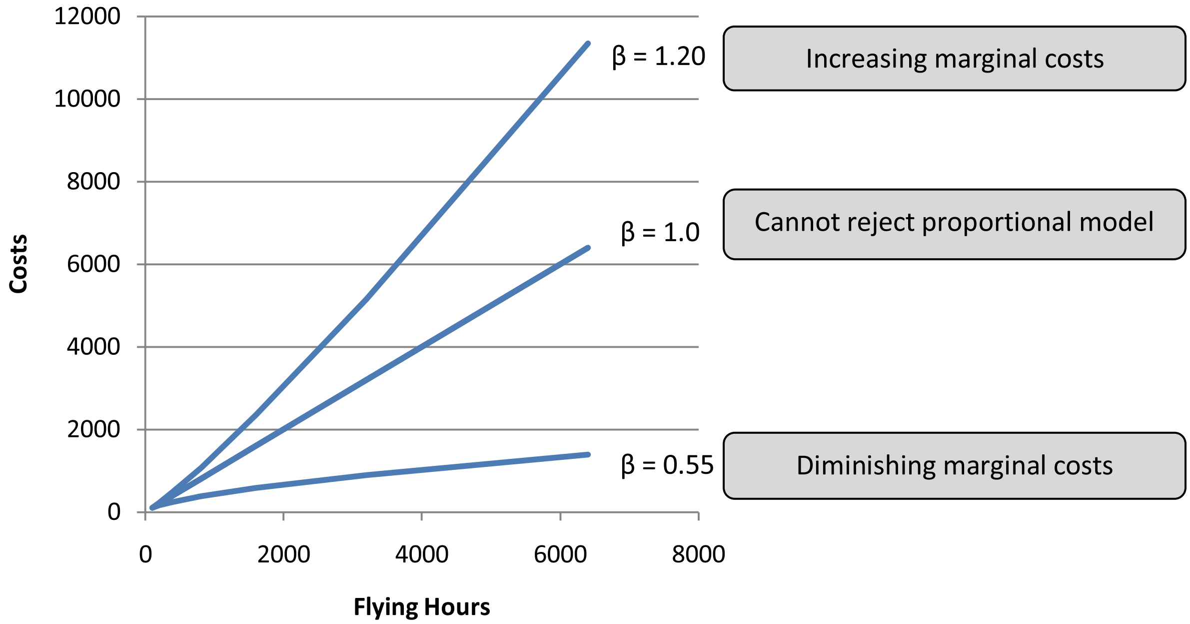
we can reject the null hypothesis that the simple regression model $\hat{\beta}_0 = \bar{y}$ is correct and the model, therefore, is significant. Since P>|t| = 0.000 for the slope of ln(flying hours), we reject the null hypothesis that the slope of the regression line equals zero; the coefficient of ln(flying hours) is statistically significant.

In the complete STATA run, several of the MD indicator variables are not statistically significantly—such as the AC-130 shown above. This result is consistent with the omitted C-130 indicator; we would not expect the estimated values for the similar AC-130 to vary greatly from the C-130 for a given amount for flying hours. We retain all of the indicators, since removing them reduces the ability to explain the coefficient on flying hours. The signs and magnitude of the indicators in the total O&S model reflect the total annual budget and age of each MD, relative to the C-130. For example, in FY06 F-16s have the largest total O&S budget of all of the MDs—$4.3 billion—and an average age of 16.9 years, compared C-130 average age of 25.7 years. The KC-135 has a $2.3 billion total O&S budget, but an average age of 45.5 years. The budget for the 20.5 year old F-117 was $341 million, less than 10% of the budget for F-16s. The small positive coefficient for the F-117 does not imply that it is a relatively inexpensive aircraft; the total annual costs for this platform are, however, less than the majority of the MDs. Although the Age appears small, it has a substantial effect on the fitted values.

What does Flying Hour Coefficient Imply?

Figure 4-3 shows the cost implications of different values of the flying hour coefficient. For constant elasticity values greater than one, we expect to see increasing marginal costs—for a doubling of flying hours, costs more than double. For values equal to one, costs increase directly with flying hours and we cannot reject the proportional model. Values in between zero and one are associated with diminishing marginal costs, which is consistent with a fixed plus variable cost model. Finally, although not depicted, a zero coefficient implies that flying hours and costs may not be related.

Figure 4-3 Implications of Flying Hour Coefficient



From Table 4-2, the coefficient of 0.56 on Ln(Flying Hours) is consistent with both nontrivial fixed costs and a marginal cost model for CPFH calculations. The constant elasticity interpretation of the log-log specification implies a possible nonlinear relationship between costs and flying hours. We will address the interpretation of the flying hour coefficient in greater detail in later chapters. For now, it is sufficient to establish a statistically significant relationship between flying hours and O&S costs.

Literature Suggests Alternative Specifications

While we have established that our estimation strategy produces a statistically significant relationship, further refinement is required to comments on CPFH metrics. Certainly the above specification is not the only possibility for testing the relationship between usage and O&S costs. The CPFH literature suggests several alternatives that we test below.

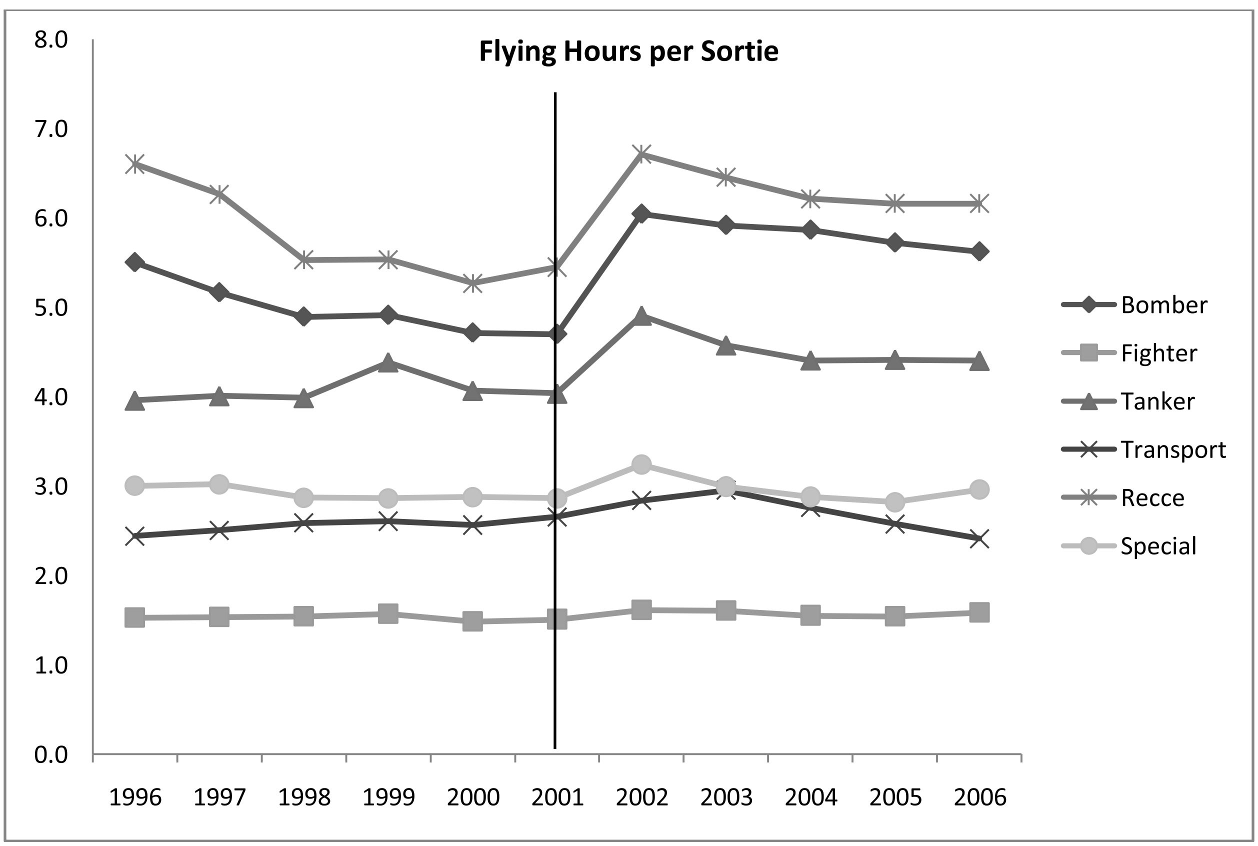
Hildebrandt and Sze (1990) controlled for the total aircraft inventory (TAI) in their examination of total O&S costs versus average flying hours. Slay (1995) found that models including number of sorties and flying hours provided better predictive power for spares costs. Hawkes (2005) and Armstrong (2006) found average sortie duration (ASD) significant in cost per flying hour (CPFH) models for F-16 and F-15, respectively.

While we found that sorties and TAI, by themselves were too highly correlated with flying hours to include in a single specification, we do not believe the ASD and TAI are too highly correlated to be included in the same model. Thus, we test the addition of ASD and flying hours per TAI.

Average Sortie Duration

The first alternative to the baseline specification we tested is the inclusion of average sortie duration. The review of the CPFH literature indicated that there were strong considerations for including sortie duration in the model specification; average sortie duration performed well in certain models, but not all. However, due to the prevalence of the average sortie duration in similar models, we investigate whether we should include it in our models. Figure 4-4 shows the flying hours per sortie or average sortie duration, with the vertical line highlighting the jump in the annual hours flown per aircraft, particularly for Transports. The chart for flying hours per landing shows a very similar pattern. [27]

Figure 4-4 Average Sortie Duration by Aircraft Type



[27] We include the figure for FH/landing in the appendix.

Although the literature specifically refers to average sortie duration, there are two related variables that might be used in this excursion: sorties and landings. A sortie is a mission that may include one or more takeoffs and landings, so sorties and landings tend to be highly correlated. It is possible that landings rather than sorties drive maintenance costs more directly, since landings are more directly related to impact wear on the aircraft. For this specification, we need to choose between the highly related sortie duration ln(FH/sortie), and flying hours per landing ln(FH/Landing).

Table 4-3 Correlations between Log-Transformed Cost, FHs , FH per Landing, FH per Sortie, and FH per TAI

	Tot Cost	FlyHrs	FH/Land	FH/TAI	FH/Sortie
Tot Cost	1.0000				
FlyingHrs	**0.7715**	1.0000			
FH/Land	0.2054	-0.2233	1.0000		
FH/TAI	-0.1633	-0.0122	0.2363	1.0000	
FH/Sortie	0.0790	-0.3058	**0.7468**	0.3661	1.0000

Table 4-3 shows the correlations at the between the different combinations of "flying Hour per" variables and costs, where the highlighted cells indicate relatively high correlations. High correlations exist between FH/Landing and FH/sortie; model selection with more than one of these parameters included may be fraught with multicollinearity issues. While landings may be a more obvious cause of wear and tear on an aircraft, many of the landings during a sortie may be partial "touch-and-go" landings. Since the impact landings may be overstated and the literature generally discusses average sortie duration, we will eliminate FH/landing in favor of FH/sortie.

Average Sortie Duration (ASD) Specification Results

The results shown in Table 4-4 present coefficients for only flying hours, age and, average sortie duration variables. The Appendix contains the results for the full models. The results for these specifications are very similar to the earlier models that did not include average sortie duration—average sortie duration is not statistically significant in the below model.

ASD Specification:
$\ln(Cost_{my}) = \beta_0 + \beta_1 * Age_{my} + \beta_2 * \ln(Flying\ Hours_{my}) + \beta_3 * LN(FH/Sortie)_{my} + \mu_m + \varepsilon_{my}$.

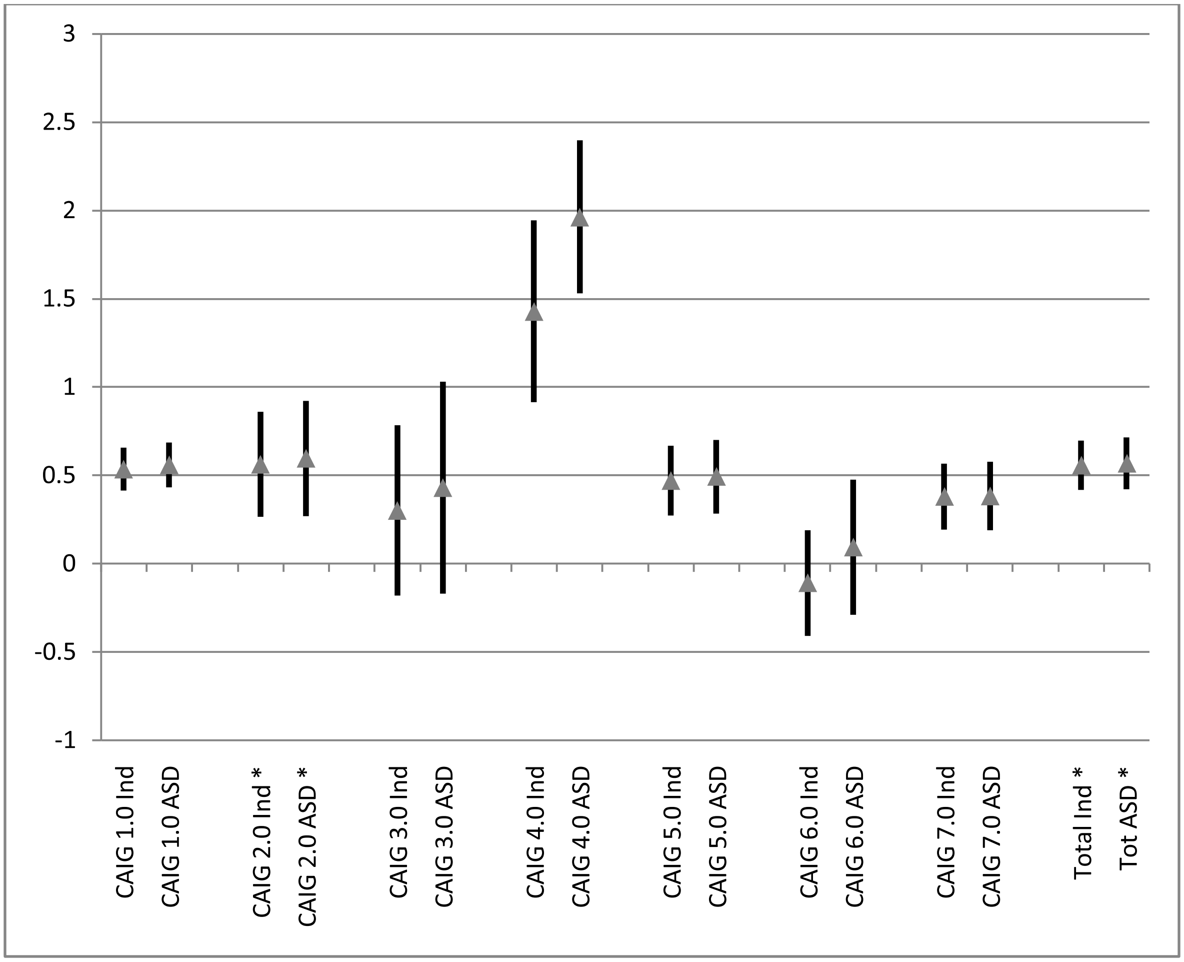
Table 4-4 ASD Specification Results

```
Linear regression                                      Number of obs =     361
                                                       F( 36,   324) = 1285.20
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9732
                                                       Root MSE      =  .25173

------------------------------------------------------------------------------
             |               Robust
lncaigtot21  |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnflyinghrs  |   .5674055   .0753162     7.53   0.000     .4192351     .715576
         age |   .0592937   .0049281    12.03   0.000     .0495987    .0689888
 lnfh_sortie |  -.2570222   .2434934    -1.06   0.292    -.7360498    .2220054
```

However, the lack of significance of ASD for total O&S costs is not sufficient to reject this specification. We also tested the inclusion of ASD at CAIG level-one to evaluate whether the variation within the disparate cost groups could be better explained with ASD. Figure 4-5 graphically shows that the inclusion of ASD does not have an appreciable effect on the coefficient estimates. This figure compares the 95% confidence intervals for the estimated flying hours coefficient, with the dark lines showing the range of the CI and the gray triangle representing the mean. Due to its lack of explanatory ability and statistical significance, we will forego the inclusion of the ASD variable in our predictive models.

Figure 4-5 Baseline versus ASD for Level-One CAIG costs

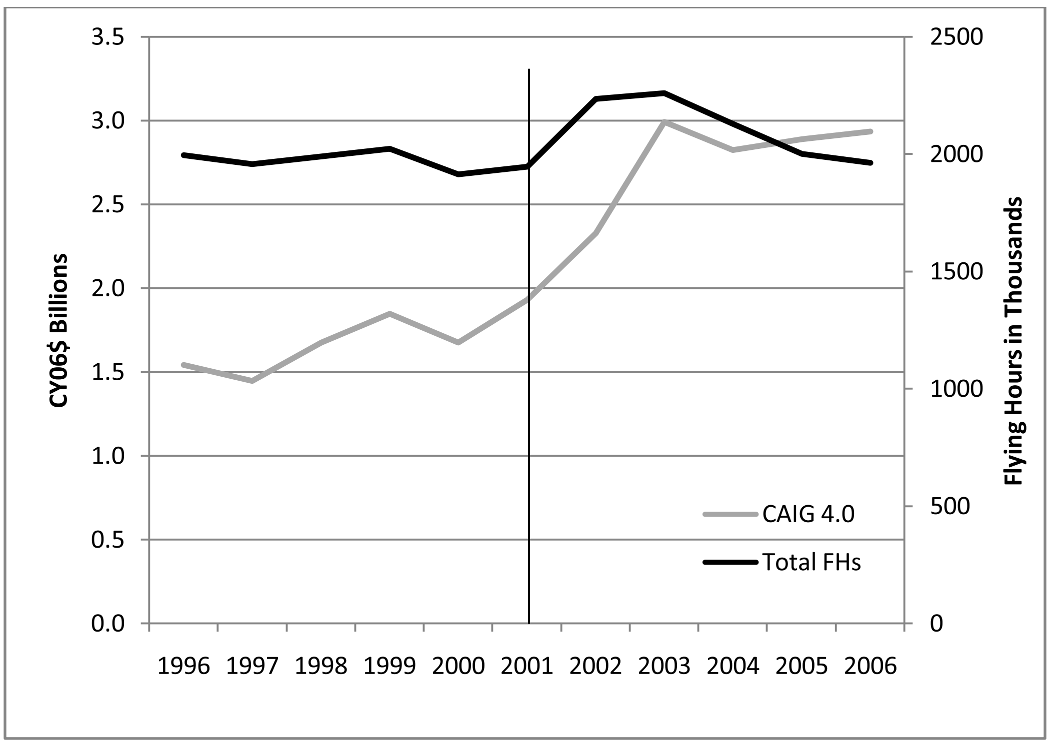


An interesting aspect to the ASD specification is that although ASD had little appreciable effect on the 95% CIs for the flying hour coefficient, it highlighted CAIG element 4.0's difference from the other element estimates. CAIG 4.0 Depot Maintenance was the only level-one category with a coefficient greater than one, so further investigation of this category is warranted.

CAIG 4.0 Deport Maintenance

The coefficient on flying hours for CAIG 4.0 is substantially different than the other CAIG elements. We suspected that the cause for this was substantially increasing costs without an analogous growth in flying hours. Figure 4-6 confirms our suspicion. The dark black line represents total flying hours for all MDs. Aside from a spike in flying hours after FY01, the overall flying hour trend is fairly flat. However, the growth in real costs for CAIG 4.0, while steady prior to FY01, substantially increases during FY02 and FY03. During FY96 to FY06, there was a significant increase in cost allocation to depot maintenance, without a commensurate increase in flying hours.

Figure 4-6 Total Flying Hours versus CAIG 4.0 Costs for All MDs



Of the 21 MDs that report CAIG 4.0 costs, 18 showed substantial growth post FY01. For example, the FY06 4.0 costs for the B-52 in were 2.4 times as large in FY06 as in FY96. Overall, there was systematic growth in CAIG 4.0 costs after FY 2001.

Figure 4-7 CAIG 4.0 Costs for six top MDs

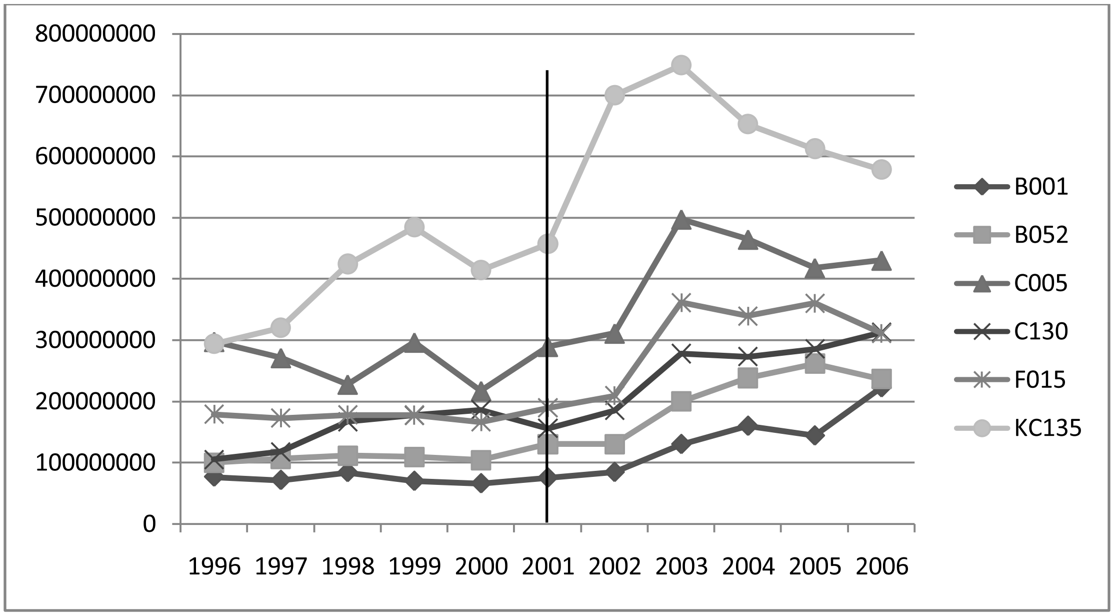
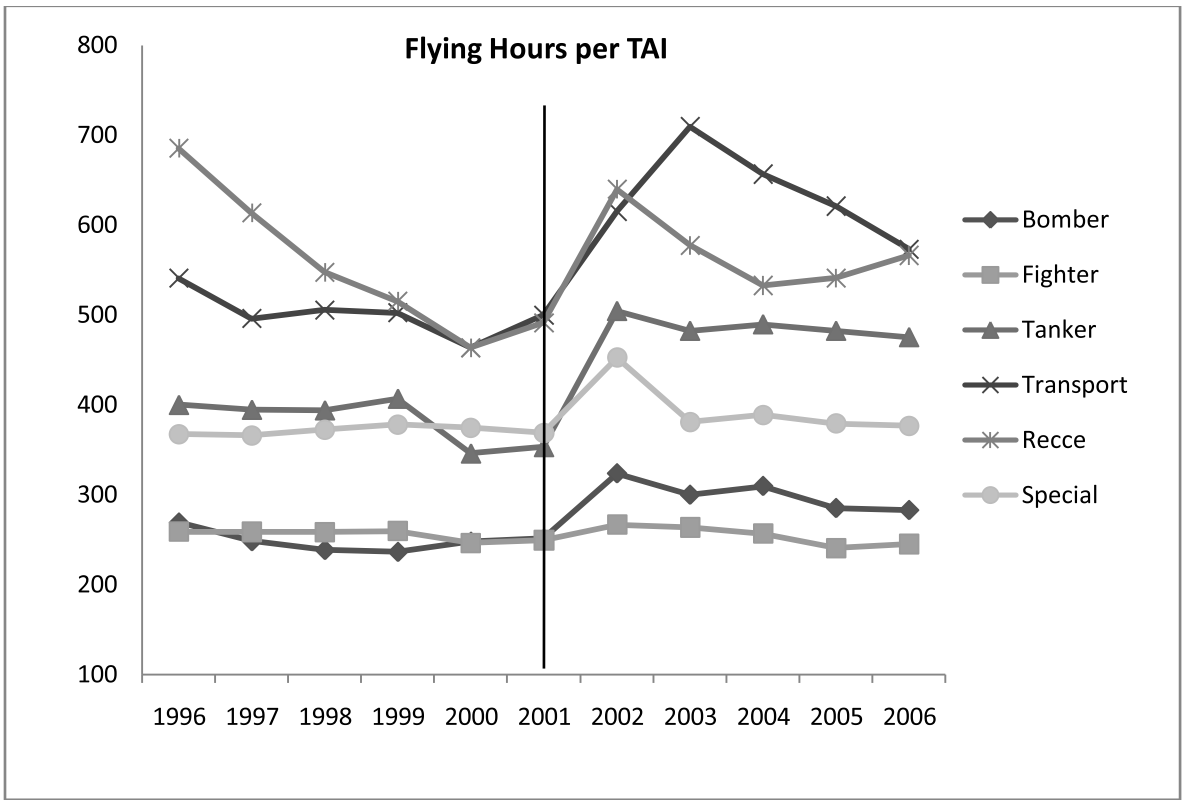


Figure 4-7 shows that two of the top six MD contributors to CAIG 4.0 costs exhibit a pattern of substantial cost growth after FY01, without related growth in flying hours. The other MDs showed large spikes in flying hours with increases in CAIG 4.0 costs occurring simultaneously or in the following fiscal year. In each case, CAIG 4.0 costs rose dramatically after FY01, irrespective of the changes in flying hours. This suggests a change in policy related to budget allocation to CAIG 4.0—a preference to spend additional dollars in depot maintenance. Our estimating approach attributes the changes in costs to flying hours; it is more likely that external policy decisions changed the relationship between depot costs and flying hours, resulting in an artificially high coefficient estimate.

Average Annual Flying Hours

Our investigation of average sortie duration highlighted another related area for investigation: average annual flying hours. Average annual flying hours is equivalent to flying hours per total aircraft inventory (TAI), for a given MD. Figure 4-8 shows that flying hours per TAI trends follow a similar pattern as those for ASD, but these two variables are not highly correlated (0.3661).

Figure 4-8 Flying Hours per Total Aircraft Inventory by Aircraft Type



Flying Hours per Total Aircraft Inventory (FH/TAI) Specification Results

The results shown in Table 4-5 present coefficients for only the flying hours, age and, FH/TAI variables.[28] The results for these specifications are again quite similar to the earlier models that did not include average sortie duration. We show this similarity in Figure 4-9. The FH/TAI variable is statistically significant in the model, but its inclusion increases the R^2 by a negligible amount, from 0. 9729 in the baseline model to 0.9745 in this model.

Indicator Specification:

$\ln(Cost_{my}) = \beta_0 + \beta_1 * Age_{my} + \beta_2 * \ln(\text{Flying Hours}_{my}) + \beta_3 * LN(FH/TAI)_{my} + \mu_m + \varepsilon_{my}$.

Table 4-5 FH/TAI Indicator Specification Results

```
Linear regression                                      Number of obs =     361
                                                       F( 36,   324) = 1445.34
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9745
                                                       Root MSE      =  .24542

------------------------------------------------------------------------------
             |               Robust
 lncaigtot21 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 lnflyinghrs |   .6788335   .0895418     7.58   0.000     .5026766    .8549903
         age |   .0611358   .0050408    12.13   0.000      .051219    .0710526
    lnfh_tai |  -.4649065   .1741936    -2.67   0.008    -.8075999   -.1222131
```

Figure 4-9 Comparison of ASD and FH/TAI Specifications to the Baseline

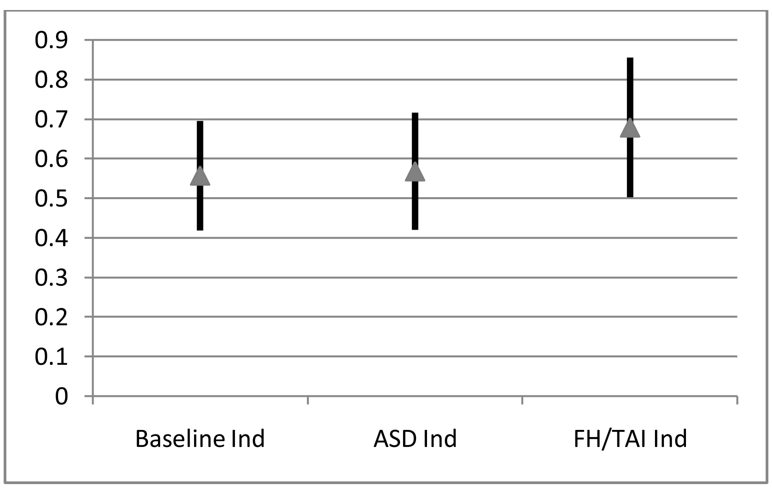


Figure 4-9 compares the 95% confidence intervals for the flying hour coefficient for total O&S costs in baseline indicator, ASD, and FH/TAI specifications. We note that the addition of the FH/TAI variable

[28] The Appendix contains the results for the full models.

moved the mean of the FH coefficient by a small amount and widened the CI. We examine more disaggregated cost models in the next chart to clarify the benefits of including the FH/TAI parameter.

Figure 4-10 CAIG Level-One Comparison of Baseline versus FH/TAI Specifications

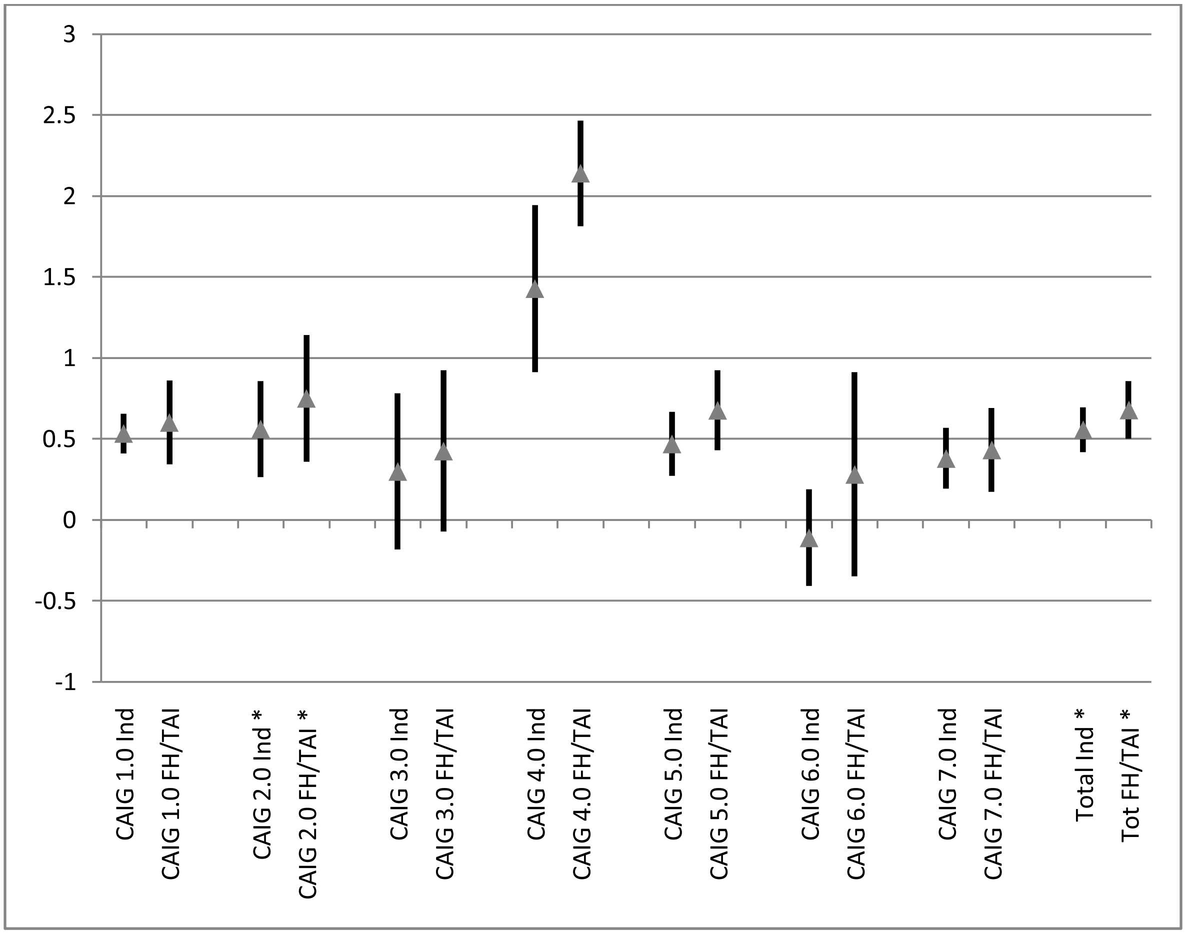


Figure 4-10 confirms the results of level-one models—FH/TAI is statistically significant for most of the level-one specifications, but its inclusion doesn't improve the statistics or interpretation. The FH/TAI specification resulted in the widening most of the 95% CIs compared to the baseline specification. Since it is unclear that the FH/TAI variable represents any kind of improvement, we will continue to use the baseline indicator specification for the remainder of the models and tests.

Performance Specification

It is a common cost estimating practice to create weapon system cost models based on performance characteristics—including aircraft weight, wingspan, thrust, and number of engines, since these variables are likely correlated with O&S costs. We examined another specification—what we term the "performance specification"—to evaluate whether performance variables can act as a proxy for the indicators. Investigating the performance specification may be a good strategy to test whether or not we can eliminate the numerous indicator variables in order to improve degrees of freedom and increase the model's ease of interpretation.

Comments on Performance Specification

We ran a nested F-test, included in the Appendix, which indicated that we do not have sufficient statistical evidence to replace the indicator variables with performance variables. The primary reason for this is that the indicator model has a very high R^2 of 0.973, compared to the performance model's 0.829. Therefore, we will continue to use the baseline indicator model as the specification for forecasting O&S budgets in Chapter Six.

Results Summary

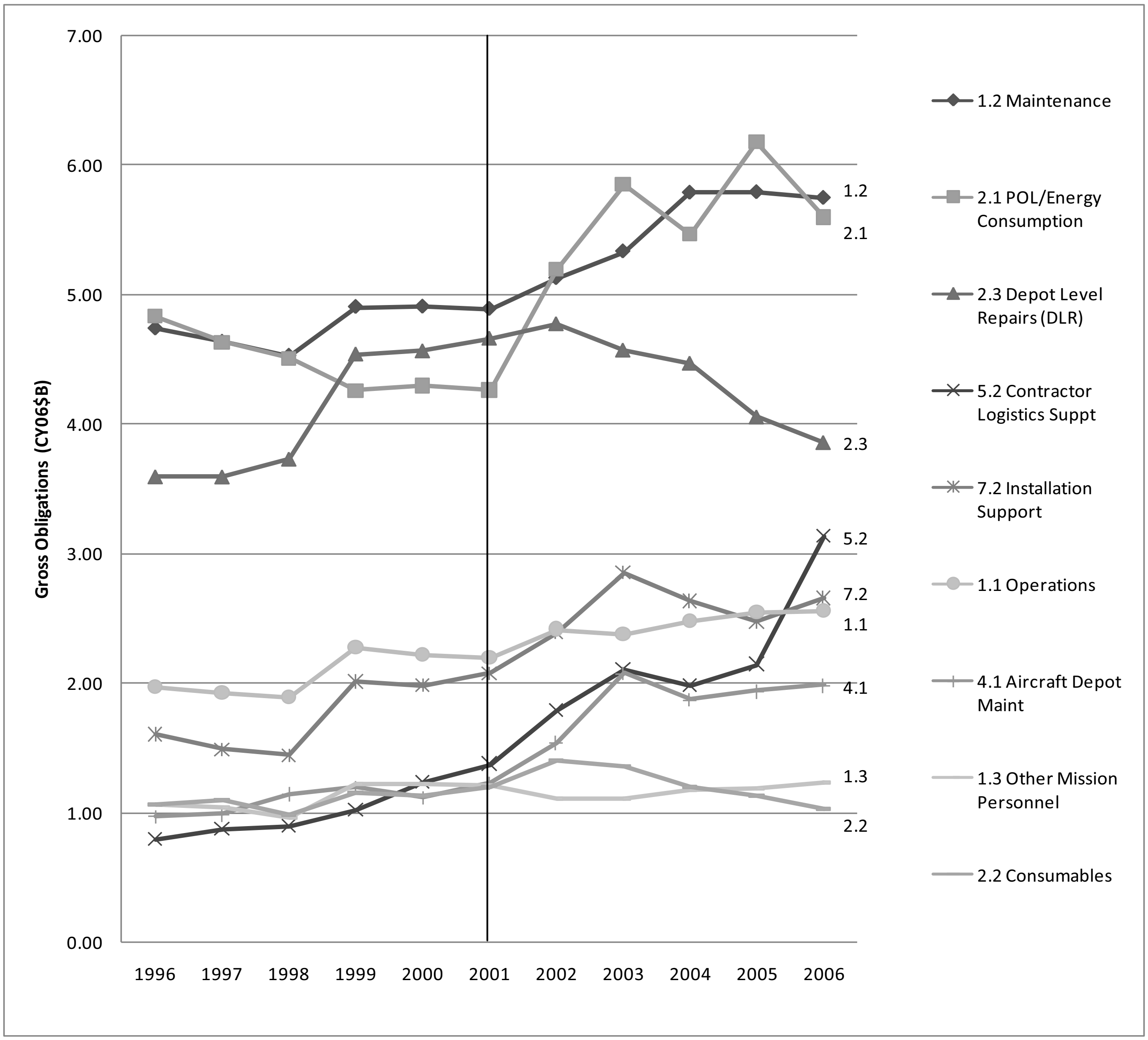
This chapter examined the relationship between flying hours and O&S costs across a number of dimensions. We created models using variations on baseline indicator specification and investigated a performance variable specification. Overall, the high degree of correlation between variables such as FHs, sorties, landings, and tails prevented them from being included in a model simultaneously. Constructed variables such as FH/sortie (ASD) were either not statistically significant—correlated with FHs—or did not contribute the explanation of cost variability. Ultimately, we retained the baseline indicator specification as the preferred specification for the investigation remaining in the next two chapters.

The constant-elasticity relationship between flying hours and costs is consistent with significant fixed costs and/or a non-linear relationship between FHs and costs. The implication of this finding includes, but is not limited to, the idea that the current Air Force CPFH factor structure should be amended to accommodate fixed and variable costs, but should also be altered to compensate for nonlinearities. The essential limitation for evaluating CPFH factors for a particular MD is the lack of data—the same reason that lead to the macro modeling of this concept. With the limitation of having only about 10 annual cost observations per MDS (or MD), the lack of statistical power of modeling may result in spurious conclusions and incorrect factors. In this chapter we have found that flying hours are useful in modeling O&S costs. Chapter Five investigates if another usage variable, Total Aircraft Inventory, is also useful in modeling O&S costs.

CAIG Element Fixed and Variable Costs

This chapter refines the analysis of Chapter Four by examining the current estimating approach for level-two CAIG costs. Figure 5-1 shows the 11-year trend for the top nine costliest level-two CAIG element costs in CY06$B—accounting for over 90% of total costs per FY. This figure shows marked increases in Element 1.2, personnel pay, Element 2.1 fuel, and element 5.2 contractor logistic support.

Figure 5-1 Top Nine CAIG Level-Two Elements by Fiscal Year (CY06$B)



The following section details the SAF/FM cost grouping of the CAIG cost elements. Of note is that the only level-two element in the top nine that shows a marked decrease after FY01 is CAIG 2.3 Depot Level Reparables.

CAIG Cost Element Grouping

Figure 5-2 SAF/FM Breakout of CAIG Level-2 Costs

- 2.1 POL/Energy Consumption
- 2.2 Consumables/Repair Parts
- 2.3 Depot Level Reparables
- 2.4 Training Munitions/Expendable Supplies
- 2.5 Other Unit Level Consumption
- 3.3 Intermediate Maintenance –Transportation
- 5.1 Interim Contractor Support
- 5.2 Contract Logistics Support
- 5.3 Other Contract Support

Variable Cost per Flying Hour

- 1.1 Operations Personnel
- 1.2 Maintenance Personnel
- 1.3 Other Mission Personnel
- 4.1 Depot Maintenance Overhaul/rework
- 4.2 Other Unit Level Consumption
- 4.3 Depot Maintenance Engine Overhaul
- 4.4 Depot Maintenance Other Equipment Overhaul

Variable Cost per TAI

- 6.1 Support Equipment Replacement
- 6.3 Other Recurring Investment
- 6.4 Sustaining Equipment Support
- 6.5 Software Maintenance Support
- 7.1 Personnel Support
- 7.2 Installation Support

Fixed Costs

Figure 5-2 shows the SAF/FM grouping of 22 OSD CAIG level-two cost elements into three categories: cost per FH, cost per TAI, and fixed costs (Lies and Klapper, 2007). The premise is that the level-two O&S costs vary proportionally with either flying hours, TAI, or are fixed. There is justification for this scheme; SAF/FM assigned the costs to these categories using corporate knowledge and reasonable assumptions about the most useful way to predict costs for each element for future budgets. However, empirical analysis will provide additional evidence as to whether the a priori categorization was proper. Our final research question addresses whether this scheme is appropriate for O&S cost forecasting and if an alternative approach might better estimate costs.

Methodology Overview

As discussed in Chapter One, SAF/FM categorized O&S CAIG cost elements into fixed, variable with flying hours, and variable with TAI. Since there is high degree of correlation between flying hours and TAI as shown in Figure 5-3, we need to test if the distinction between costs that vary with flying hours and costs that vary with TAI produces a better predictive relationship for level-two costs. We will examine the a priori level-two categorization by running two similar, but distinct indicator model specifications: one with flying hours as the variable of interest and the other with TAI as the variable of interest.

The first part of this analysis is to establish whether the coefficients on FH or TAI are statistically significant. For a given level-two CAIG element, if the coefficient on the variable of interest is statistically significant, it indicates that there is empirical evidence to support a correlation with that variable and costs. The converse of that statement is also true. The statistical significance of the coefficients of interest will show which CAIG elements relate to FHs, TAI, or both. Due to the high correlation between FHs and TAI, it is plausible that those coefficients will be statistically significant in the same specifications—where only those variables change.

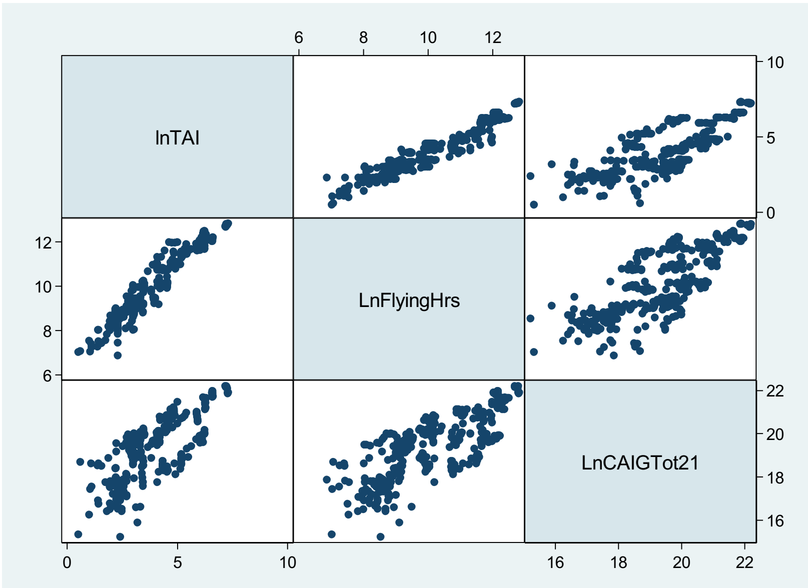
Another component of this analysis is the comparison of the coefficients. Again, given the high correlation between FHs and TAI, it is likely that they will both relate to costs in a similar way. We will establish which specification provides better estimates or narrower confidence intervals.

Evaluating Variable and Fixed Costs

As described above, SAF/FM groups OSD CAIG level-two costs elements into three categories: cost per FH, cost per TAI and fixed costs. In this section, we test to see whether these groups of costs are appropriate.

There is strong a priori evidence to suggest that the distinction between costs that vary with flying hours and costs that vary with TAI is artificial. Due to the very strong correlation between flying hours and TAI, it is likely that either variable would provide similar explanatory capability vis-à-vis costs.

Figure 5-3 Correlations Between LN(TAI), LN(FHs) and Ln(costs)



The high correlation of flying hours and TAI—correlation coefficient of 0.9564—shown in Figure 5-3 makes this an interesting test to perform. The initial naïve model of including both flying hour and TAI variables in the same specification is not sound since multicollinearity causes substantial problems with model selection. Instead, we create a flying hour specification and a TAI specification to evaluate which of the variables (or both) has a statistically significant relationship with a particular subset of O&S costs.[29] We run models for two similar specifications, one for each of the level-one and level-two disaggregated CAIG costs:

FH specification: $\ln(\text{Cost}_{my}) = \alpha + \beta^*\text{Age}_{my} + \delta^*\ln(\text{Flying Hours}_{my}) + \mu_m + \varepsilon_{my}$

TAI specification: $\ln(\text{Cost}_{my}) = \alpha + \beta^*\text{Age}_{my} + \delta^*\ln(\text{Total Aircraft Inventory}_{my}) + \mu_m + \varepsilon_{my}$

According to the above argument, the flying hour coefficient estimates of these two specifications models should be very similar. We expect that the coefficient on Ln(FH) and Ln(TAI) should both be significantly different than one; coefficients not statistically different than one would be consistent with a proportional model. They should also be similar in magnitude for a given cost element. If the 95% confidence interval (CI) for either coefficient includes zero, it supports the idea that O&S costs do not vary with that coefficient of interest, for that element.

[29] This approach does not address the difficult question of which specification is better, since a comparison of R^2 does not provide that insight.

Of note is that given SAF/FM assumptions, Element 2.0, 3.0, and 5.0 costs should vary with flying hours; Element 1.0 and 4.0 costs should vary with total aircraft inventory (TAI), and Element 6.0 and 7.0 costs should vary with neither FH nor TAI.

Figure 5-4 SAF/FM Categorization of CAIG Level-one elements

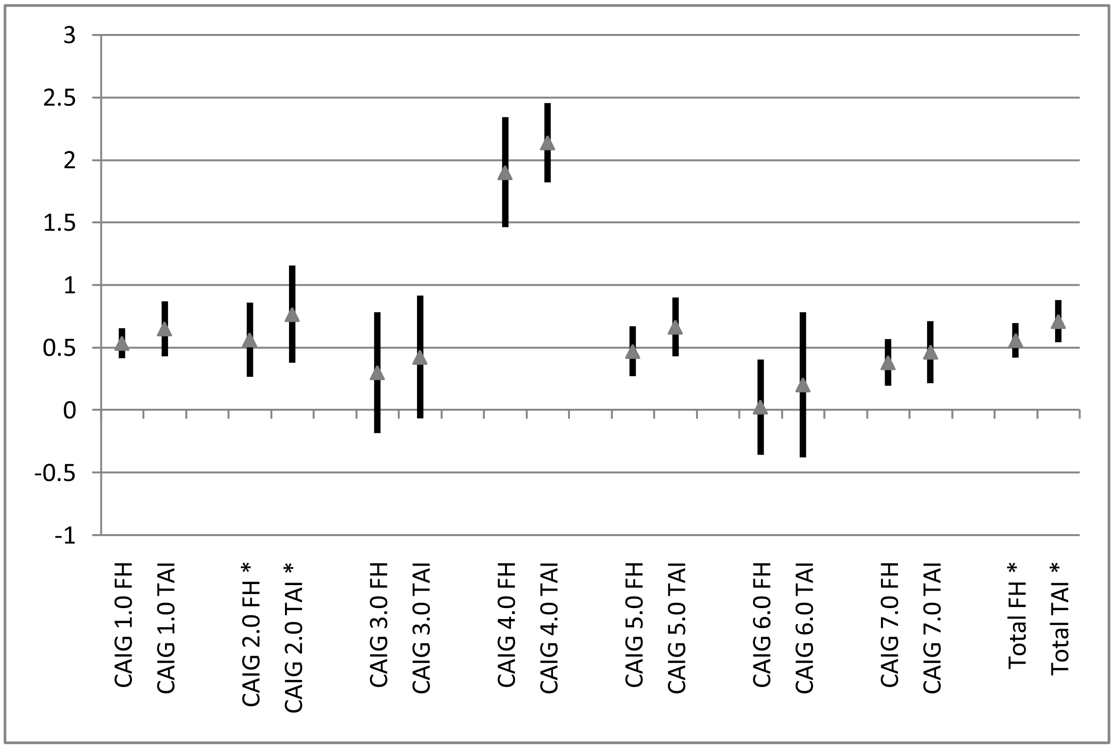
Variable with FH	**Variable with TAI**	**Fixed Costs**
CAIG 2.0 CAIG 3.0 CAIG 5.0	CAIG 1.0 CAIG 4.0	CAIG 6.0 CAIG 7.0

Figure 5-4 shows the ex ante SAF/FM Categorization of CAIG cost elements. Each of the level-two elements falls into the same group as the level-one, so this diagram is aligns with the level-two depiction, just at a higher level of aggregation. We used the baseline indicator specification to test SAF/FM's cost grouping.

Using the baseline indicator specification with FHs and TAI separately, we evaluated twenty-eight separate of cost models—one for each of the level-one, level-two, and total CAIG categories. Figure 5-5 depicts the results for the level-one and the total cost aggregations. The general finding that the 95% CIs overlap for the two specifications shown at level-one is also valid at level-two. Due to the complexity of the figure, however, we show only the results for level-one.

Figure 5-5 FH versus TAI specifications for Level-one CAIG Categories

**Does not include 2.1 Fuel costs*



There are several interesting results in Figure 5-5. First, it appears that the FH and TAI specifications find very similar results. Secondly, it appears that CAIG Elements 3.0 Intermediate Maintenance and 6.0 Sustaining Support do not vary with either FHs or TAI. This supports the SAF/FM categorization of element 6.0, but contradicts it for element 3.0. Finally, the lower bound 95% CIs for element 4.0 are both greater than one. This implies that as either FHs or TAI double, costs for that element more than double, consistent with highly variable costs. A possible interpretation for this phenomenon is that, given depot constraints, increasing demand for repairs has driven up costs.[30] Another possibility is that this is a spurious relationship, given that depot demand is calendar driven, rather than flying hour driven. [31]

One of the problems with the above approach is the flying hours and TAI cannot be included in the same model due to multicollinearity. In order to more directly address the effect of flying hours on costs, while controlling for TAI, we examine a normalized cost specification in the next section.

---

[30] Element 4.0 is a special case, in that we know that the AFCAIG inflation index is lower vis-à-vis the inflation actually experienced in this area. We examine the effect of a different inflation correction in the Appendix.
[31] Depot visits occur after a set amount of time—five to six years—not the amount of time the aircraft has flown.

Normalized Cost Model

A possible remedy to the problem of including both flying hours and TAI in the same model is to normalize the dependent cost variable and the primary independent usage variable with TAI. In other words, divide the dependent cost variable by TAI and divide the independent flying hour variable by TAI. The basic models is: $\ln(Cost/TAI)_{my} = \alpha + \beta{*}Age_{my} + \delta{*}\ln(Flying\ Hours/TAI)_{my} + \mu_m + \varepsilon_{my.}$ This allows testing of whether a given OSD CAIG element is sensitive to flying hours; if a particular element is sensitive to flying hours, the coefficient on LnFH/TAI would be significant. In this specification, the constant term is of important, since it will give an indication of the amount of fixed costs

Figure 5-6 Normalized Costs Results

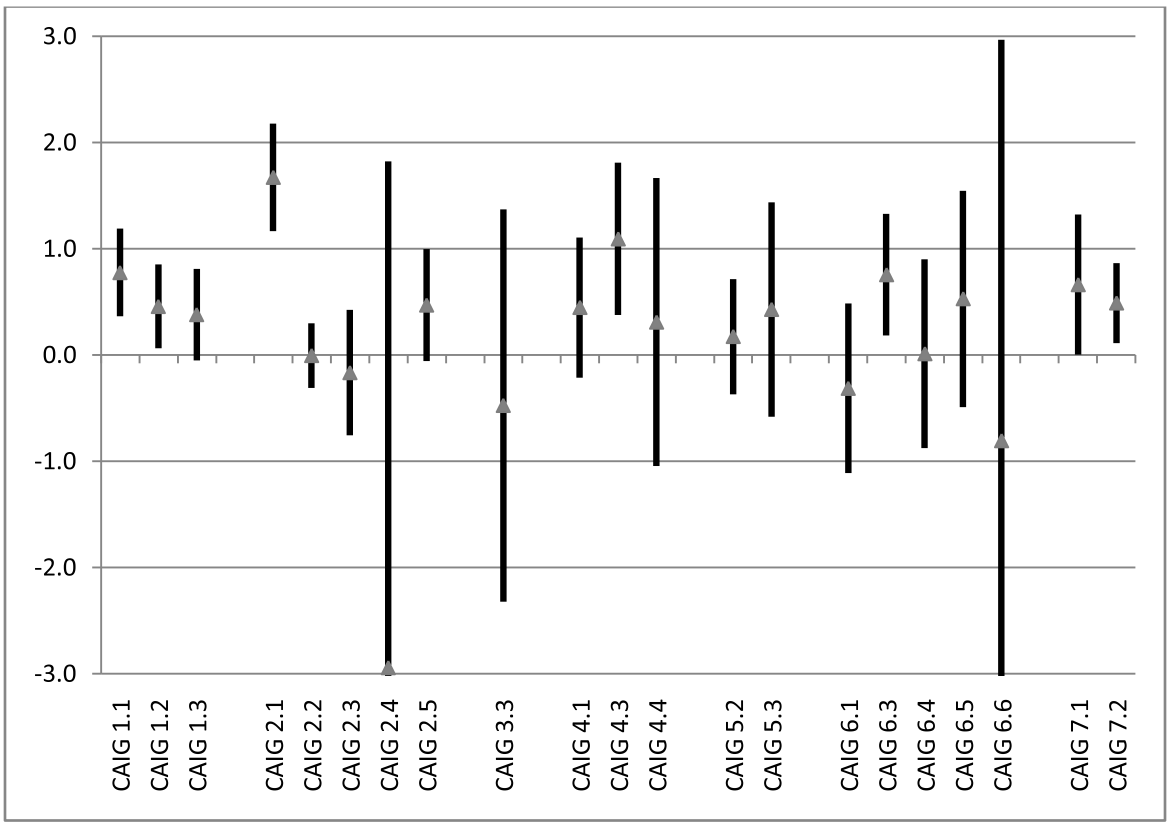
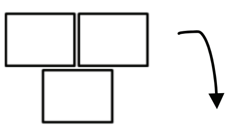


Figure 5-6 provides slightly more refined results: elements 2.1, 4.3, 6.3, and 7.2 appear to be sensitive to flying hours. In general, this supports the findings from the earlier tests in this chapter. It is also consistent with the expectations that fuel, element 2.1, would be highly sensitive to variations in flying hours. While SAF/FM characterizes element 6.0 and 7.0 as fixed costs, each of the specifications in this chapter indicate that certain costs in these elements are sensitive to flying hours.

Updated CAIG level-two Categorization

The results above show that there does not appear to be a forecasting advantage in using TAI versus FH categories—both parameters provide essentially the same information in the log-log specification, due to their highly correlated nature. Although there is a high degree of correlation between flying hours and TAI, it makes more sense to use flying hours in predictive models due to the lack of variation in TAI within and between fiscal years. We apply the logic from Figure 4-3, then, to evaluate the degree of fixed costs present in the CAIG level-two elements. Instead of the ex ante SAF/FM categorization, we propose the following structure for the CAIG level-two elements, which ranges from substantially fixed with respect to flying hours to highly variable with respect to flying hours:

Figure 5-7 Updated CAIG Structure



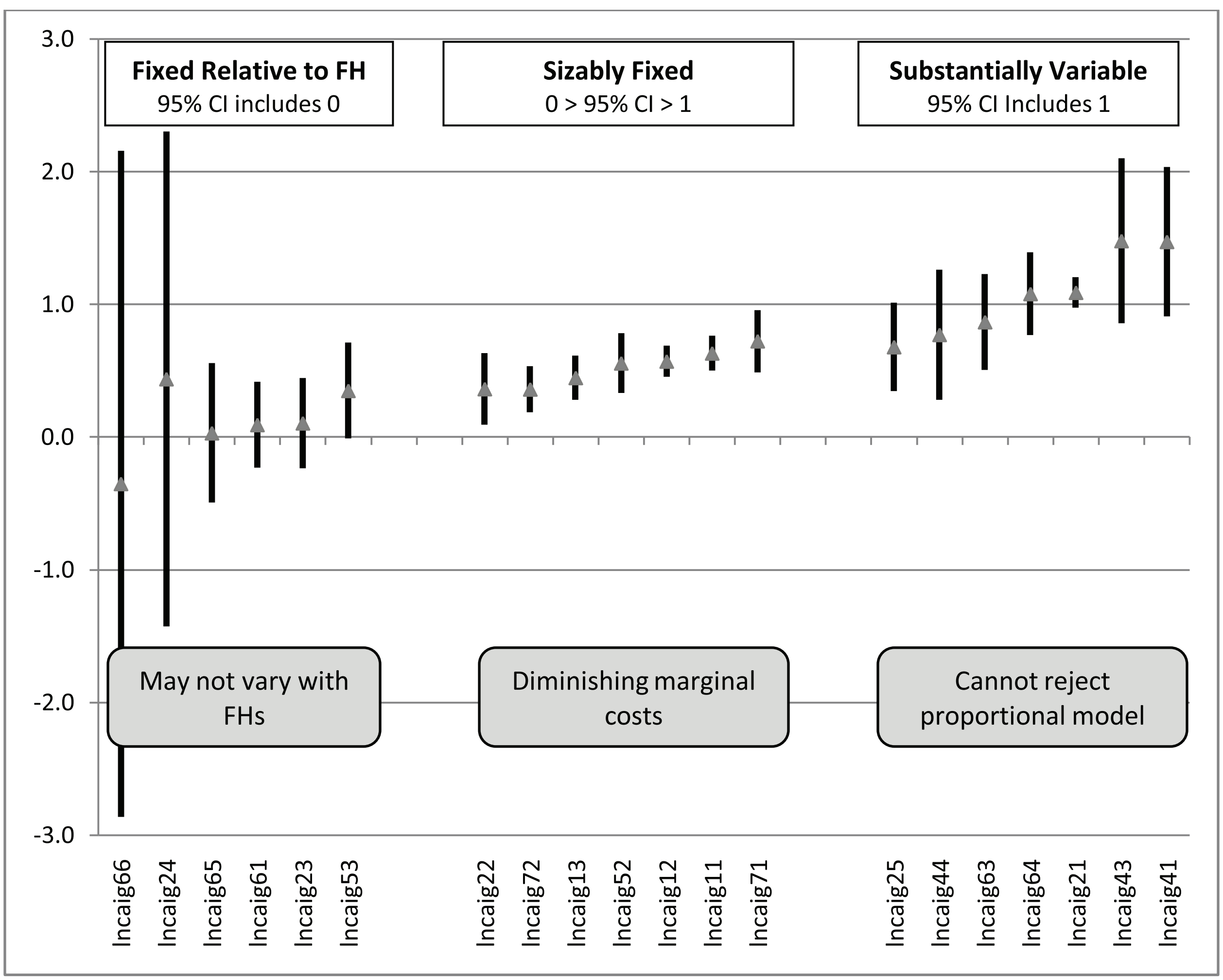


Figure 5-7 shows the 95% CIs for the level-two CAIG costs elements. The dark black line indicates the range of the confidence interval, while the gray triangle represents the mean. This analysis indicates a change in the SAF/FM cost element grouping; we shift from SAF/FM's scheme that delineates between costs that vary with flying hours and costs that vary with TAI to one that evaluates cost variation only with flying hours.

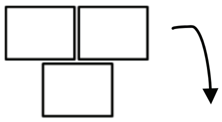
Figure 5-7 shows our proposed grouping of CAIG level-two elements: "Fixed Relative to Flying Hours," "Sizably Fixed," and "Substantially Variable." The figure shows that some of the level-two elements may not vary with FHs—those whose CI includes zero. We designate those elements "Fixed Relative to Flying Hours." Since the CIs for these elements include zero, we do not have sufficient statistical evidence to reject a slope of zero. Although, our grouping discusses the elements in this group as being fixed relative to flying hours, the high degree of correlation between flying hours and TAI makes it likely that these costs would be fixed relative to TAI also.

A second group, whose CI lies between zero and one—indicative of diminishing marginal costs—we define as "Sizably Fixed;" the costs in elements vary with flying hours but contain a nontrivial amount of fixed costs. Our definition of "Sizably Fixed" implies only that the model is consistent with a fixed and variable cost model, with greater fixed costs associated with smaller non-negative coefficient (less than one, but greater than zero) on the FHs variable.

The third group is comprised of elements that are consistent with costs that are substantial variable with flying hours—those whose 95% CIs include values greater than one. Since each of the 95% CIs for the coefficient estimates in the last group includes one, we cannot reject the proportional model for them. Overall, the new grouping indicates shift to additional fixed costs. We interpret costs in this category as those that would more than double with the doubling of flying hours. Cost elements described as "Substantially Variable" may include a fixed cost component.

An important feature of the results from Figures 5-6 and 5-7 is that they validate our known prior information about CAIG 2.1 fuel costs. We know that POL costs are approximately proportionally related to flying hours. Our results for 2.1 show not only that coefficient in approximately equal to one, but that the 95% confidence interval is quite narrow. In figure 5-6, CAIG 2.1 is clearly the element most sensitive to flying hours. The estimated coefficient of 1.089 is consistent with the ex ante assumption that fuel cost are proportional to flying hours and that the proportional model works well here—precisely the reason why we excluded these costs for the aggregated models.

Figure 5-8 Proposed Change in Categorization of CAIG Elements



Variable with FH	Variable with TAI	Fixed Costs
All sub-elements of CAIG 2.0, 3.0, 5.0	All sub-elements of CAIG 1.0, 4.0	All sub-elements of CAIG 6.0, 7.0

Substantially Variable	Sizably Fixed	Fixed Relative to FH
CAIG 2.1	CAIG 1.1	**CAIG 2.3**
CAIG 2.5	CAIG 1.2	**CAIG 2.4**
CAIG 4.1	CAIG 1.3	**CAIG 5.3**
CAIG 4.3	**CAIG 2.2**	CAIG 6.1
CAIG 4.4	**CAIG 5.2**	CAIG 6.5
CAIG 6.3	CAIG 7.1	CAIG 6.6
CAIG 6.4	CAIG 7.2	

Figure 5-8 shows our proposed changes to the current SAF/FM categorization of CAIG cost elements, based on our results from Figure 5-7. One item to note is that the SAF/FM categories are not named in exactly same way as our suggested categories. SAF/FM delineates variable costs between costs that vary with flying hours and costs that vary with TAI. In the SAF/FM scheme, fixed costs refer to costs that do not vary with either flying hours or TAI. Our empirical evidence suggests that costs vary with both FHs and TAI in approximately the same way for log-log specifications. With the exception of 2.1 fuel costs, there is little substantial statistical evidence to support the hypothesis that FHs explains variation in costs any differently than TAI. Changes to the SAF/FM categorization are bolded. We noted that several elements changed from "not related" to variable costs, including: sustaining support costs in 6.3, 6.4, and indirect support costs in 7.1, 7.2. Also, several elements changed from variable to "not-related," including: depot level reparables 2.3, 2.4, and contractor support costs in 5.3. CAIG element 5.3 is a borderline categorization and may belong in the Sizably Fixed category.

In general, the empirical evidence supports a change the SAF/FM categorization. We believe that their delineation between costs that vary with flying hours and costs that vary with TAI is largely artificial in predictive models. Our categorization shows a substantial shift from variable costs to fixed costs. Our analysis shows that CAIG elements 2.2, 2.3, 2.4, 5.2, and 5.3—representing approximately \$8.4B of the \$30.5B FY06 O&S budget—shifted from a SAF/FM variable category to one of our fixed categories. While not all of the costs in these elements are fixed, the presence of nontrivial fixed costs implies that the proportional metric would misestimate costs. CAIG elements 6.3 and 6.4 shifted to our "Substantially Variable" category, but total only \$0.28B in FY06. Our most substantive critique of the SAF/FM approach is that it underplays the presence of fixed costs, reducing the effectiveness of a proportional cost forecasting method. We will address the budget implications of our updated cost categorization the next chapter.

CHAPTER SIX

Cost per Flying Hour and Marginal Costs

The goal of our analysis is to improve Air Force resource allocation by enhancing its ability to forecast costs. The critical empirical analysis of usage effects helps inform the Cost per Flying Hour calculation, providing a basis to examine the larger policy picture. This chapter addresses the implications of a marginal CPFH metric vis-à-vis the current proportional metric.

What does the coefficient on Flying Hours imply?

The preceding chapters analyzed whether O&S cost varied with flying hours and whether flying hours or TAI was the appropriate primary dependent variable. Based on the results, we believe that the baseline indicator specification, $\ln(\text{Cost}_{my}) = \alpha + \beta^*\text{Age}_{my} + \delta^*\ln(\text{Flying Hours}_{my}) + \mu_m + \varepsilon_{my}$, remains preferable to its alternatives. We will use the results of the baseline indicator specification to examine the implications of marginal costs to the current Air Force CPFH metric.

We showed in Chapter Four that the coefficient on the Ln(flying hours) variable is 0.5567, for the baseline indicator specification. The constant elasticity interpretation of the log-log model implies a nonlinear relationship between costs and FH. The general economic interpretation for this model specification is diminishing marginal returns; as flying hours increase, costs increase at a decreasing rate. Although the coefficient of 0.56 on Ln(Flying Hours) implies such a nonlinear relationship when transformed from log-space back to levels, it is consistent with nontrivial fixed costs and a marginal cost model for CPFH.

Figure 6-1 Total Annual O&S Costs by Flying hours for the B-2

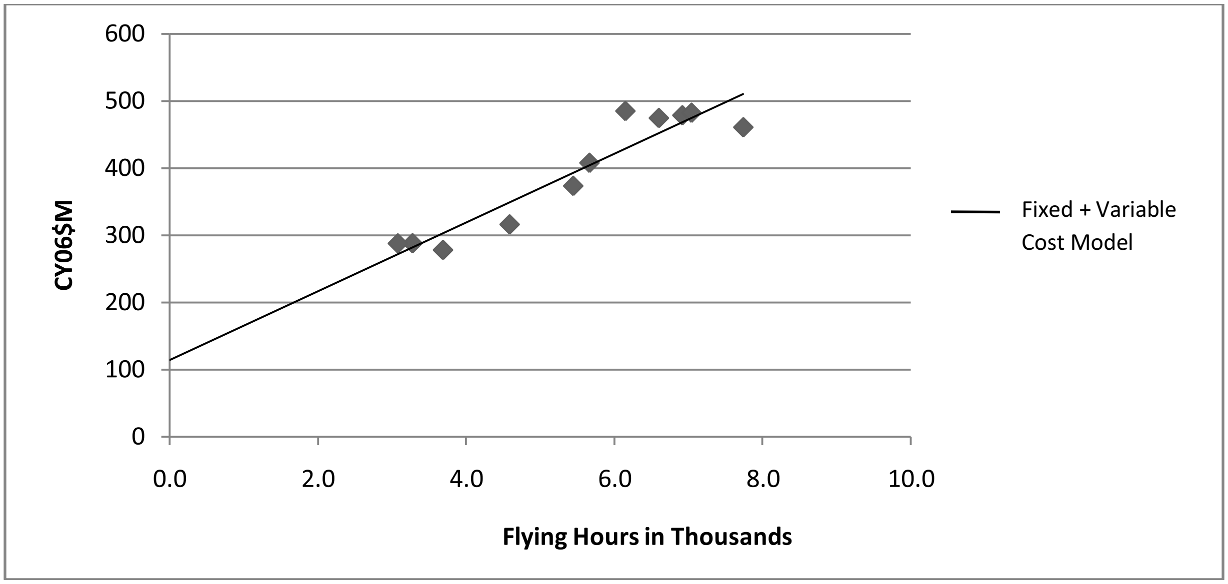
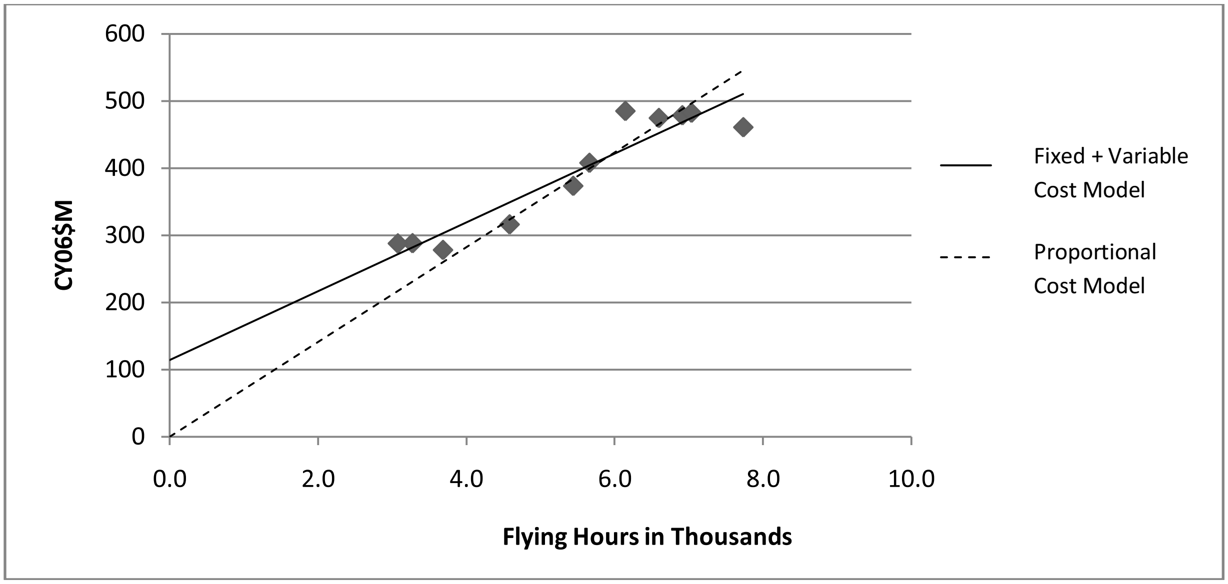


Figure 6-1 depicts the association between flying hours and annual costs in base year dollars for the B-2 Spirit; the solid line represents the single MD regression line for the B-2. For this platform, it seems reasonable that the CPFH should include fixed and variable cost components. Other Air Force platforms show a similar relationship between flying hours and costs—O&S costs should be broken into fixed and variable components.

However, the current CPFH metric dictates a significantly different estimation form. For a given year and MDS, the CPFH metric divides the historical costs by the number of flying hours. This procedure is analogous to a line that goes through the origin and implies that costs double as flying hours double. Figure 6-2 shows a comparison of the of the current CPFH metric—the proportional model—to our suggested specification. The dashed line shows the results of the proportional model, while the unbroken line represents a model with both fixed and variable costs. These two models will predict costs differently, particularly at the extreme points of the data range.

Figure 6-2 Current CPFH Metric for B-2 Costs



Is the resulting 0.56 coefficient on FH high or low?

In general, the body of literature generally supports a Fixed plus Variable CPFH model. Armstrong (2006), Hawkes (2005) , and Hildebrandt (1990) all found empirical evidence that indicates a CPFH factor specification should include both a fixed and variable component to improve the estimation of O&S costs. However, neither Armstrong nor Hawkes' results can be compared directly to this study. Hildebrandt finds a coefficient of 0.614 on Ln(average flying hours).[32] Although this specification differs to the one used in this study, it does generally support the idea of nontrivial fixed costs and, possibly, a marginal cost model for CPFH.

[32] The basic model specification is Ln(Costs / TAI) = Ln(FHs / TAI) + Ln(TAI) + Ln(flyaway costs) +age. This is a substantially different specification than the one tested in this section of the study. However, we test a specification similar to Hildebrandt's model in Chapter Four.

Cost per Flying Hour and Marginal Costs

Based on the current construction of Cost per Flying Hour (CPFH) factors, as aircraft fly more we expect their maintenance costs to increase linearly; USAF uses CPFH to determine O&M budgets by multiplying projected flying hours against the CPFH factor. Superficially, it seems reasonable that if we fly an aircraft twice as far, O&M costs double. However, the initial results from the baseline modeling of this study show that the doubling of flying hours actually increases maintenance costs by about fifty-six percent. This implies that CPFH factors, even those with both a fixed and variable component, may misestimate maintenance costs associated with aircraft usage. If costs increase 56% as flying hours double, the average cost CPFH metric overestimates positive changes in flying hours. Similarly, decreasing flying hours using the current CPFH metric reduces costs too much—it underestimates costs when the flying hour program is reduced[33]. Moreover, given the constant elasticity interpretation of the log-log specification, the relationship between flying hours and maintenance costs is potentially nonlinear—further reducing the accuracy of CPFH metrics.

Since SAF/FM, the organization responsible for creating and amending AF budgets, routinely use the proportional cost per flying hour (CPFH) metric as a marginal cost (SAF/FMC, 2005), we will investigate whether this results in over-budgeting. For additional flying hours, the AF simply multiplies the CPFH factor by the number of projected hours required. However, the CPFH factor is the average cost per flying hour by fleet, so may be misleading when used as a marginal cost. For example, during the course of a fiscal year, planners may realize that they require additional flying hours over the original estimate for a particular platform. Using the CPFH, which is calculated as an average cost, to estimate the cost of additional hours likely results in an overestimation of required funding. Breaking costs into fixed and variable components may provide a more accurate method for calculating maintenance budgets. This increased accuracy will also reduce the potential for distorting the overall O&M budget allocation.

[33] The current CPFH factors may be adequate for calculating initial budgets. However, estimating O&M cost changes to those initial budgets with an average costs metric causes problems.

While previous chapters included all of the O&S costs less fuel in the analysis, this section purposefully will focus on O&M costs. We confine this portion of the analysis to a subset of O&M costs—CAIG 2.0—since they are the costs associated with the flying hour program and CPFH metrics.

Table 6-1 Indicator Specification for CAIG 2.0 Costs

```
Linear regression                                      Number of obs =     361
                                                       F( 35,   325) = 2595.07
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.9861
                                                       Root MSE      =  .23023

------------------------------------------------------------------------------
             |               Robust
    lncaig20 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 lnflyinghrs |   .8875841    .055778    15.91   0.000     .7778526    .9973156
         age |   .0282624   .0039167     7.22   0.000     .0205571    .0359676
```

Using the indicator baseline specification, $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + \mu_m + \varepsilon_{my}$, we examine the level-two CAIG 2.0 costs in more detail. Despite the continued use of the proportional CPFH model employed by the Air Force, the empirical evidence shows that the proportional model may not apply for certain cost categories. The results from the CAIG 2.0 indicator specification in Table 6-1 indicate that the doubling of flying hours actually increases maintenance costs by about eight-nine percent. However, examining the disaggregated costs shows a more complex story. Since CAIG Element 2.0 comprises the costs typically considered in CPFH calculation—consumable supplies, spare parts, and aviation fuel—we examine each of the sub-elements separately.

While we would prefer to estimate costs using the fixed + variable cost model shown in Figures 6-1 and 6-2, sample size problems severely limit the validity of that estimation. Instead, we estimate a common elasticity across MDs for each of the level-two CAIG 2.0 costs. We will then use these common elasticities to demonstrate the potentially substantial differences between the proportional CPFH metric and the common elasticity metric. It is important to note that the common elasticity models behave similarly to the linear fixed+variable cost models across the range of flying hour values within particular MDs; both models accommodate the presence of fixed costs and tend to dampen cost changes associated with changes to flying hours compared to the proportional model.

Table 6-2 CAIG Level-Two Indicator Specification Results

CAIG Element	ln(flyingHrs) Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
2.0 Total	0.888	0.0558	15.91	0.000	0.7779	0.9973
2.1 POL/Energy Consumption	1.089	0.0593	18.36	0.000	0.9723	1.2057
2.2 Consumables	0.362	0.1362	2.66	0.008	0.0941	0.6300
2.3 Depot Level Reparables (DLR)	0.104	0.1729	0.60	0.547	-0.2361	0.4445
2.4 Training Munitions	0.437	0.9397	0.47	0.643	-1.4287	2.3029
2.5 Other Unit Level Consumption	0.679	0.1691	4.01	0.000	0.3459	1.0113

Table 6-2 provides rather mixed results for the relationship between flying hours and the various 2.0 sub-elements. For 2.1 AVFUEL, the relationship is statistically significant. Since the CAIG 2.1 95%CI includes one, we cannot reject the proportional model—as anticipated. Neither 2.3 nor 2.4, by virtue of the 95% CI including zero, shows a statistically significant relationship with flying hours.[34] However, 2.2 and 2.5 both including consumable items, show the anticipated relationship: the coefficient on flying hours is statistically significant and less than one.

The level-one element 2.0 costs appear to support the overall hypothesis that as flying hours increase, costs increase at a lesser rate. However, the disaggregated view shows that model 2.1 Fuel costs is consistent with the proportional model, but that other elements, such as 2.2 consumables, likely should follow a fixed plus variable cost model. Using the coefficients for the level-two categories for CAIG 2.0, we examine the budget implications of changing from the current proportional model to a constant elasticity model in the next section.

[34] Element 2.4 may suffer from a sample size issue, with only 112 observations and is excluded

Budget Implications of a Marginal Cost per Flying Hour

In this section, we test to see how changes in the CPFH computation affect future budgets. We have access to both future flying hours and future budgets in CAIG format (FY07 through FY11) in the Air Force Cost and Performance (AFCAP) Tool. Focusing on CAIG 2.0 costs, AFCAP shows a reduction in flying hour requirements during in both FY07 and FY08. Although AFCAP does provide flying hour and cost information for fiscal years beyond FY08, these data are straight-lined and do not provide insight into the different cost forecasting methods. AFCAP does not provide future budget information about element 2.4 Training Munitions.

Table 6-3 Comparison of Forecasting Methods (CY06$)[35]

	2006	**2007**	**2008**
Flying Hours (FH)	2,113,643	1,904,229	1,887,797
Proportional Method Forecast			
CY 2.1 - POL/Energy Consumption	$5,733,658,248	$5,165,583,716	$5,121,008,787
CY 2.2 - Consumables	$1,131,641,875	$1,019,522,021	$1,010,724,347
CY 2.3 - DLR	$4,100,710,903	$3,694,424,144	$3,662,544,166
CY 2.5 - Other Unit Level Consumption	$545,131,973	$491,121,849	$486,883,854
Proportional Total	$11,511,143,000	$10,370,651,730	$10,281,161,154
Constant Elasticity Method Forecast			
CY 2.1 - POL/Energy Consumption	$5,733,658,248	$5,115,025,083	$5,066,958,096
CY 2.2 - Consumables	$1,131,641,875	$1,091,054,488	$1,087,646,279
CY 2.3 - DLR	$4,100,710,903	$4,058,457,080	$4,054,814,865
CY 2.5 - Other Unit Level Consumption	$545,131,973	$508,459,099	$505,479,917
Constant Elasticity Total	$11,511,143,000	$10,772,995,750	$10,714,899,157
Delta		**$402,344,020**	**$433,738,002**

Table 6-3 shows an aggregated cost forecast, using the SAF/FM standard proportional model and the constant elasticity model supported by this study's analysis. We predicted CAIG level-two costs for element 2.0, using the change in annual flying hours, in this example. While this particular "top-down" method would typically not be used to construct annual budgets, it is useful in demonstrating the

[35] The coefficient on flying hours variable for CAIG 2.3 is not statistically significant. We include it here to demonstrate the differences between the forecasting methods more clearly.

differences between the two specifications.[36] Using the constant elasticity model to predict costs differed from the proportional model: CAIG 2.1 decreased by $51M or 0.8%, CAIG 2.2 increased by $72M or 5.8%, CAIG 2.3 increased by $367M or 8.2% and CAIG 2.5 increased by $17M or 2.9%. In aggregate, the difference in predicted costs between the proportional method and the constant elasticity method can amount to hundreds of millions of dollars—a substantial amount in an $11B budget.

AFI 65-503 states that the proportional CPFH factors are used to create the budgets and also make changes to them during the fiscal year. Without a variable and fixed component, this appears to be an incorrect approach. Table 6-3 shows that a change in the forecasting method has budget implications, where flying hours decrease. The constant elasticity forecast was calculated using the elasticities shown in Table 6-2, and we constructed the proportional method forecast with elasticities equal to one. Table 6-3 is consistent with the assertion that it is possible that the proportional model reduces the budgets too far, when flying hours decrease. Similar logic applies to flying hour changes within a given fiscal year. AFI 65-503 explicitly states that the Air Force uses proportional CPFH factors "to build as well as increment and decrement" requirements based on flying hours. Unless all of the costs are variable, then the current CPFH metric misestimates budgets when flying hour requirements change.

Results Summary

This chapter examines the budget implications of the relationship between flying hours. In determining future budgets, the constant elasticity cost model can produce substantially different results than the current proportional model. The distinction between these two models becomes particularly important when making incremental changes to budgets during a fiscal year; the proportional model includes fixed costs in the incremental changes that do not belong there. In an environment where we anticipate changes to flying hour profiles, it is critical to accommodate the marginal cost model to prevent misallocation of resources.

[36] While either the proportional model or the constant elasticity model will provide usable results for annual forecasts, the proportional model exaggerates costs for incremental changes within a fiscal year. The constant elasticity method is better Incremental flying hour change

CHAPTER SEVEN
Conclusions

Operating and Support costs are a vital and substantial component of the total Air Force Budget. Since O&S includes both Operations and Maintenance (O&M) and military personnel (MILPERS) appropriations, it funds the vast majority "must-pay" bills—virtually all funding that keeps the aircraft in the air falls under O&S. Therefore, improving the Air Force's ability to forecast O&S costs is critical for ensuring that sufficient resources are available to accomplish Air Force missions.

We found that total O&S costs vary with flying hours; we examined multiple different specifications to estimate a common cost elasticity across MDs. Our baseline indicator specification showed that as flying hours double, total O&S costs increase by fifty-six percent. We found that adding average sortie duration or average annual flying hours to the baseline indicator specification did not enhance our ability to the relationship between O&S costs and flying hours.

Subsequently, we noted that flying hours and TAI are highly correlated, such that either is acceptable when predicting O&S costs; models that use either as the primary independent variable will produce similar results in log-space. However, flying hours provide more direct insight into potential variability in mission requirements. The static nature of TAI implies the need for other covariates to accurately predict costs. While we found substantial evidence that the 2.1 fuel costs vary with flying hours, the results are less clear for the other elements. We believe that the current SAF/FM categorization emphasizes an artificial distinction between "variable with TAI" and "variable with flying hours." We established an alternative scheme which places increased emphasis on fixed costs; our categories are "Fixed Relative to Flying Hours," "Sizably Fixed," and "Substantially Variable." The presence of nontrivial fixed costs in our categorization implies that the proportional metric would misestimate costs.

USAF may experience a drawdown in terms of both personnel and aircraft. Reducing the number of aircraft will affect sortie durations, total flying hours, and flying hours per aircraft. Using improperly specified Cost per Flying Hour factors could adversely affect flying hour budgets and the Air Force's ability to employ air power. In the context of a drawdown, using the current proportional model may reduce budgets too far. To mitigate this potential budget forecast problem, our analysis indicates that the CPFH metric should include fixed and variable costs. Moreover, the current CPFH metric has some intrinsic flaws—not only in its calculation, but in its application. AFI 65-503 states that many of the specific CPFH factors can be used to calculate both initial budgets and incremental changes to budgets. Unless these factors include only variable costs, then the factors will overestimate or underestimate costs for an increase or decrease in flying hours, respectively. Our empirical evidence supports element 2.1 fuel as an inherently variable cost (with FHs), but other elements should not be calculated with a proportional model.

Finally, the paucity of cost data is the root cause of many of the limitations of this and other analyses. We encounter substantial sample size problems when modeling system-specific costs and interpretation problem for pooled models. The choice between creating models by MDS with 11 FH-allocated cost observations and creating a model for a collection of dissimilar aircraft at an aggregated level to navigate around the sample size problem is not enviable. However, the Air Force can improve cost data collection and reporting to enhance its ability to forecast costs. While it is unlikely that the Air Force will collect tail-level cost data due to costs and limitations of the current accounting systems, it is reasonable to expect that they could increase the frequency of reporting. In fact, AFTOC began reporting cost data quarterly in FY06—a substantial improvement from a sample size perspective. Ideally, the Air Force could report O&S cost data monthly, similar to the monthly cost reports required of many research and development programs. While this imposes additional data costs, it would result in better statistical models used to forecast a $40 billion annual budget.

Appendix

CAIG Cost Element Descriptions

The "Operating and Support Cost-Estimating Guide," published by The Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG) provides definitions for the level 1 and level 2 CAIG cost elements. The following are summaries of the level 1 elements:

1.0 MISSION PERSONNEL The mission personnel element includes the cost of pay and allowances of officer, enlisted, and civilian personnel required to operate, maintain, and support operational systems. This includes the personnel necessary to meet combat readiness, training, and admin requirements.

2.0 UNIT-LEVEL CONSUMPTION Unit-level consumption includes the cost of fuel and energy resources; operations, maintenance, and support materials consumed at the unit level; stock fund reimbursements for depot-level reparables; operational munitions expended in training; transportation in support of deployed unit training; temporary duty pay; and other unit-level consumption costs, such as equipment leases.

3.0 INTERMEDIATE MAINTENANCE Intermediate maintenance performed **external to a unit** includes the cost of labor and materials and other costs expended by designated activities/units in support of a primary system and associated support equipment. Intermediate maintenance activities include calibration, repair, and replacement of parts, components, or assemblies, and technical assistance.

4.0 DEPOT MAINTENANCE Depot maintenance includes the cost of labor, material, and overhead incurred in performing major overhauls or maintenance on a defense system, its components, and associated support equipment at centralized repair depots, contractor repair facilities, or on site by depot teams. Some depot maintenance activities occur at intervals ranging from several months to several years.

5.0 CONTRACTOR SUPPORT Contractor support includes the cost of contractor labor, materials, and overhead incurred in providing all or part of the logistics support to a weapon system, subsystem, or associated support equipment. The maintenance is performed by commercial organizations using contractor or government material, equipment, and facilities.

6.0 SUSTAINING SUPPORT Sustaining support includes the cost of replacement support equipment, modification kits, sustaining engineering, software maintenance support, and simulator operations provided for a defense system. War readiness materiel is specifically excluded.

7.0 INDIRECT SUPPORT Indirect support includes the costs of personnel support for specialty training, permanent changes of station, and medical care. Indirect support also includes the costs of relevant host installation services, such as base operating support and real property maintenance.

CAIG Level 2 Element Descriptions

The following are summaries for the nine highest dollar value CAIG level-two elements, as shown in Figure 5-1:

1.1 OPERATIONS. The pay and allowances for the crew or full complement of personnel required to operate a system, including officers and enlisted personnel.

1.2 MAINTENANCE. The pay and allowances of military and civilian personnel who support and perform maintenance on a primary system, associated support equipment, and unit-level training devices. Depending on the maintenance concept and organizational structure, this element will include maintenance personnel at the organizational level and possibly the intermediate-level. Organizational maintenance personnel normally perform on-equipment maintenance; intermediate maintenance personnel perform off-equipment maintenance.

1.3 OTHER MISSION PERSONNEL. The pay and allowances of military and civilian personnel who perform unit staff, security, or other mission support functions.

2.1 POL/ENERGY CONSUMPTION. The unit-level cost of petroleum, oil, and lubricants (POL); propulsion fuel; and fuel additives. May also include field-generated electricity and commercial electricity necessary to support the operation of a system.

2.2 CONSUMABLE MATERIAL/REPAIR PARTS. The costs of material consumed in the operation, maintenance, and support of a primary system and associated support equipment at the unit level. Examples of **maintenance** material in this category include consumables and repair parts such as transistors, capacitors, gaskets, fuses, and other bit-and-piece material. Examples of **non-maintenance** material include coolants, deicing fluids, tires, filters, batteries, paper, diskettes, ribbons, and maps.

2.3 DEPOT-LEVEL REPARABLES. The unit-level cost of reimbursing the stock fund for purchases of depot-level reparable (DLR) spares used to replace initial stocks. DLRs may include repairable individual parts, assemblies, or subassemblies that are required on a recurring basis for the repair of major end items.

4.1 Aircraft Depot Maintenance. The labor, material, and overhead costs for overhaul or rework of aircraft returned to a centralized depot facility. Includes programmed depot maintenance, analytic condition inspections, and unscheduled depot maintenance.

5.2 CONTRACTOR LOGISTICS SUPPORT. Contractor logistics support (CLS) includes the burdened cost of contract labor, material, and assets used in providing support to a weapon system, subsystem, and associated support equipment. **CLS funding covers depot maintenance** and, as negotiated with the operating command, necessary organizational and intermediate maintenance activities.

7.2 INSTALLATION SUPPORT. Consists of personnel normally assigned to a host installation who are required for the unit to perform its mission in peacetime. Functions performed by installation support personnel include costs for personnel pay and materials to support to system-specific mission personnel. Base operating support activities may include communications, supply operations, personnel services, installation security, base transportation, and real property maintenance.

Table A-1 shows the unabridged level-one and level-two CAIG costs from FY96 through FY06 from AFTOC.

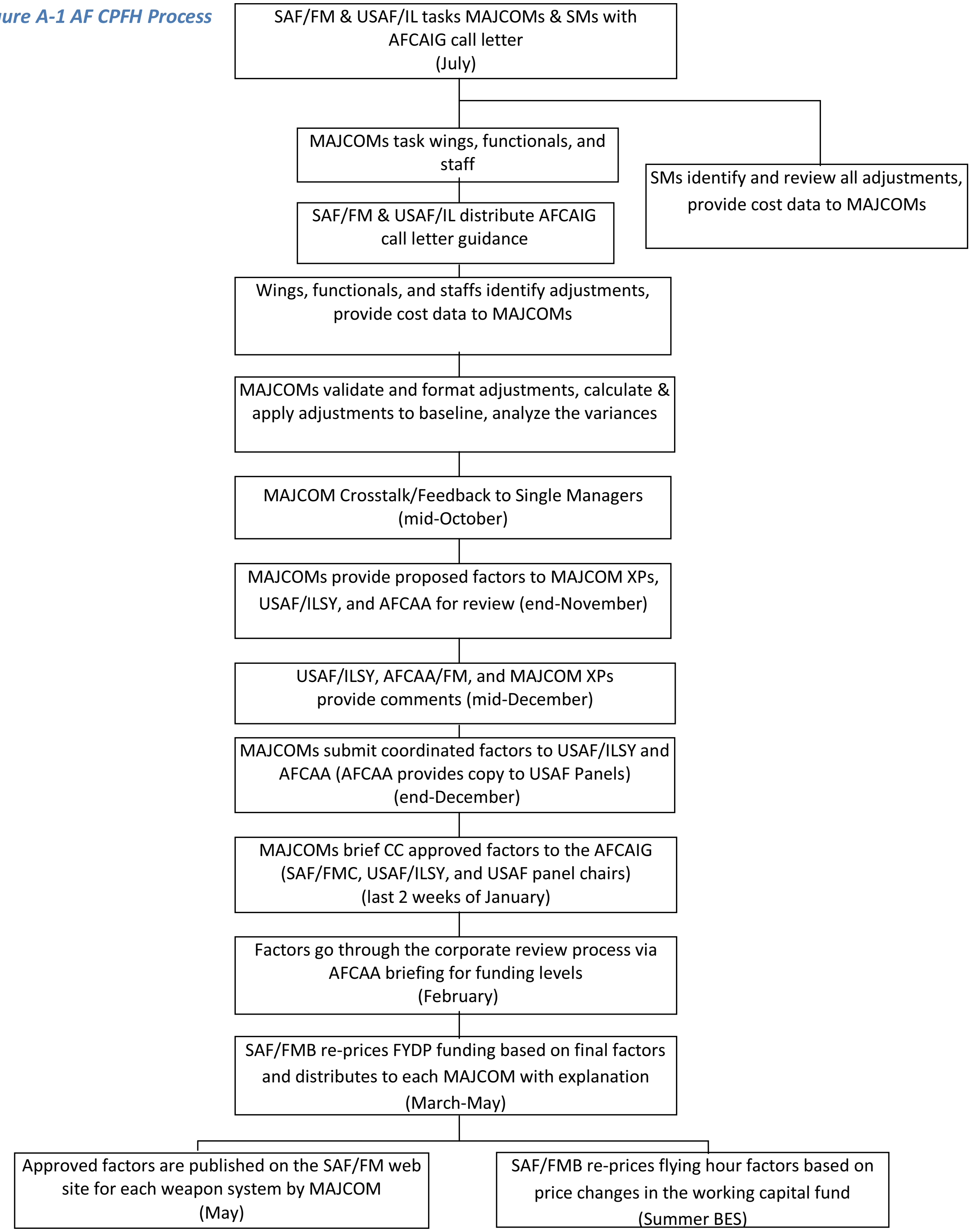
Table A-1 CAIG O&S Cost Breakout for FY96-FY06 (Billions $CY06)

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	FY96-FY06
1.0 Mission Personnel	**7.77**	**7.59**	**7.37**	**8.39**	**8.34**	**8.29**	**8.64**	**8.81**	**9.44**	**9.52**	**9.53**	**93.72**
1.1 Operations	1.97	1.92	1.89	2.27	2.22	2.20	2.42	2.38	2.48	2.55	2.56	24.85
1.2 Maintenance	4.74	4.64	4.53	4.91	4.91	4.89	5.13	5.34	5.79	5.79	5.75	56.40
1.3 Other Mission Personnel	1.06	1.04	0.96	1.21	1.21	1.21	1.10	1.10	1.17	1.18	1.23	12.46
2.0 Unit-Level Consumption	**10.04**	**9.77**	**10.00**	**10.68**	**10.68**	**10.80**	**12.28**	**12.45**	**11.90**	**12.04**	**11.30**	**121.96**
2.0 w/o 2.1 Fuel	**5.21**	**5.14**	**5.49**	**6.42**	**6.38**	**6.53**	**7.09**	**6.60**	**6.43**	**5.86**	**5.71**	**66.87**
2.1 POL/Energy Consumption	4.83	4.63	4.51	4.26	4.30	4.26	5.19	5.85	5.47	6.18	5.60	55.09
2.2 Consumables	1.06	1.10	0.98	1.15	1.13	1.20	1.40	1.36	1.20	1.13	1.03	12.72
2.3 Depot Level Repairs (DLR)	3.60	3.60	3.73	4.54	4.57	4.66	4.78	4.57	4.47	4.06	3.86	46.45
2.4 Training Munitions			0.34	0.27	0.25	0.25	0.27	0.25	0.30	0.24	0.35	2.53
2.5 Other Unit Level Cons	0.55	0.45	0.44	0.46	0.43	0.42	0.65	0.42	0.46	0.43	0.47	5.18
3.0 Intermediate Maintenance	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**
4.0 Depot Maintenance	**1.54**	**1.45**	**1.68**	**1.85**	**1.67**	**1.93**	**2.33**	**2.99**	**2.82**	**2.88**	**2.93**	**24.07**
4.1 Aircraft Depot Maint	0.97	0.99	1.14	1.20	1.12	1.22	1.54	2.07	1.87	1.94	1.98	16.03
4.3 Engine Depot Maint	0.56	0.44	0.49	0.57	0.50	0.62	0.68	0.75	0.73	0.70	0.66	6.70
4.4 Other Depot Maint	0.02	0.02	0.04	0.08	0.05	0.08	0.11	0.16	0.21	0.25	0.30	1.33
5.0 Contractor Support	**0.87**	**0.93**	**0.96**	**1.08**	**1.30**	**1.44**	**1.86**	**2.18**	**2.06**	**2.21**	**3.24**	**18.12**
5.2 Contractor Logistics Suppt	0.79	0.87	0.90	1.01	1.23	1.37	1.79	2.10	1.98	2.14	3.13	17.32
5.3 Other Contractor Suppt	0.07	0.06	0.07	0.07	0.06	0.07	0.07	0.08	0.08	0.07	0.11	0.80
6.0 Sustaining Support	**0.66**	**0.55**	**0.58**	**0.63**	**0.61**	**0.65**	**0.82**	**0.84**	**0.67**	**0.49**	**0.43**	**6.95**
6.1 Support Equip Replacement	0.11	0.08	0.07	0.07	0.05	0.04	0.08	0.14	0.07	0.03	0.03	0.76
6.3 Other Recurring Investment	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.11
6.4 Sustaining Engineering Supt	0.35	0.25	0.24	0.26	0.25	0.26	0.36	0.38	0.31	0.31	0.27	3.23
6.5 Software Maintenance	0.17	0.20	0.24	0.28	0.28	0.33	0.36	0.28	0.27	0.12	0.10	2.63
6.6 Simulator Operations	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.04	0.02	0.02	0.02	0.22
7.0 Indirect Support	**1.81**	**1.71**	**1.70**	**2.30**	**2.29**	**2.37**	**2.76**	**3.23**	**3.01**	**2.85**	**3.03**	**27.06**
7.1 Personnel Support	0.21	0.22	0.25	0.29	0.30	0.30	0.37	0.37	0.38	0.37	0.38	3.44
7.2 Installation Support	1.61	1.49	1.44	2.01	1.98	2.08	2.39	2.85	2.64	2.48	2.66	23.62
CAIG Total	**22.70**	**22.01**	**22.29**	**24.94**	**24.89**	**25.49**	**28.70**	**30.50**	**29.90**	**30.00**	**30.47**	**291.88**
CAIG Total w/o 2.1 Fuel	**17.87**	**17.37**	**17.78**	**20.68**	**20.59**	**21.22**	**23.51**	**24.65**	**24.43**	**23.82**	**24.87**	**236.79**

Air Force CPFH Process

Figure A-1 shows the process by which the Air Force approves CPFH factors (AFCAIG, 1999).

Figure A-1 AF CPFH Process



Regression Results

MD Indicators with Age—Baseline Model

Table A-2 shows the results for the baseline indicator specification for Total CAIG costs (w/o CAIG 2.1): $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + \mu_m + \varepsilon_{my}$.

Table A-2 Baseline Indicator Model Results

Linear regression

Number of obs	= 361
F(35, 325)	= 1278.66
Prob > F	= 0.0000
R-squared	= 0.9729
Root MSE	= .25232

lncaigtot21	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnflyinghrs	**.5567066**	**.0701463**	**7.94**	**0.000**	**.4187086**	**.6947046**
age	.0580409	.0047241	12.29	0.000	.0487472	.0673347
a010	-.0904037	.0609626	-1.48	0.139	-.2103349	.0295275
ac130	-.187737	.2321092	-0.81	0.419	-.6443632	.2688892
at038	-2.467883	.2125825	-11.61	0.000	-2.886095	-2.049672
b001	1.020294	.1768703	5.77	0.000	.6723389	1.368249
b002	1.368357	.2882702	4.75	0.000	.8012463	1.935469
b052	-.8257119	.1757006	-4.70	0.000	-1.171366	-.4800578
c005	.2488935	.1030718	2.41	0.016	.0461213	.4516657
c009	-1.47033	.1895251	-7.76	0.000	-1.843181	-1.097479
c017	.3135769	.1821264	1.72	0.086	-.0447186	.6718724
c020	-1.083791	.2733818	-3.96	0.000	-1.621612	-.5459694
c021	-1.648628	.1259146	-13.09	0.000	-1.896338	-1.400917
c026	-1.668475	.3286681	-5.08	0.000	-2.315061	-1.02189
c037	-.6309188	.3148186	-2.00	0.046	-1.250258	-.0115793
c141	-.9522639	.1348017	-7.06	0.000	-1.217458	-.6870698
e003	.0210957	.1748094	0.12	0.904	-.3228051	.3649965
e008	1.059993	.2940038	3.61	0.000	.4816027	1.638384
ec130	-2.402496	.3059647	-7.85	0.000	-3.004418	-1.800575
f015	.8814558	.0529971	16.63	0.000	.7771952	.9857164
f016	1.030994	.0676762	15.23	0.000	.8978556	1.164133
f117	.2649531	.2199092	1.20	0.229	-.1676722	.6975783
hc130	-1.363564	.2334061	-5.84	0.000	-1.822741	-.9043861
kc010	-.489251	.1148488	-4.26	0.000	-.7151919	-.26331
kc135	-.945577	.0893918	-10.58	0.000	-1.121437	-.7697174
lc130	-1.10233	.3028709	-3.64	0.000	-1.698165	-.5064948
mc130	-.5495226	.1602679	-3.43	0.001	-.864816	-.2342293
rc135	-1.078702	.2209355	-4.88	0.000	-1.513346	-.6440574
t001	-1.221405	.1134917	-10.76	0.000	-1.444676	-.9981337
t006	-1.066625	.1889684	-5.64	0.000	-1.438381	-.6948694
t037	-2.369316	.0929886	-25.48	0.000	-2.552251	-2.18638
t038	-1.887523	.075945	-24.85	0.000	-2.036929	-1.738117
t043	-2.660843	.2886502	-9.22	0.000	-3.228702	-2.092985
u002	.293296	.2364258	1.24	0.216	-.1718221	.7584142
wc130	-2.306326	.3133242	-7.36	0.000	-2.922725	-1.689926
_cons	13.2254	.8848913	14.95	0.000	11.48456	14.96624

Performance Specification

To better align our analysis with common cost estimating practices, we constructed a specification that uses airplane dimension and performance characteristics to predict cost. This specification performed well statistically, but was rejected by an F-test in favor of the baseline indicator model.

Aircraft Performance Data

Military aircraft performance data are available from a wide variety of sources, none of which is consistent in the fields they report or, indeed, in the values for given fields. Therefore, to completely fill in the data matrix for performance data, it is necessary to consult multiple sources. We used a combination of the AF Factsheets, Jane's All the World's Aircraft, and Periscope databases to complete the matrix.

The aircraft in the USAF inventory break into two main categories, in terms of performance characteristics: jets and props. An interesting note on the comparison of jet aircraft to propeller driven aircraft is that there is no common measure of power: jets use "thrust" and props use "shaft horsepower (SHP)." We wanted to avoid the modeling issue where the thrust variable perfectly correlates with jets and the SHP variable perfectly correlates with props. There is a complex conversion from SHP to thrust, however we use a very simplified conversion. We convert SHP to thrust using the rule of thumb: Thrust /(2.5 x 0.8) = Equivalent SHP, where the 0.8 refers to the propeller efficiency. From this approximation, we simply multiply SHP by two to arrive at thrust for props. Thrust and SHP are reported per engine, so multiplying the new equivalent thrust variable by the number of engines gives total thrust.

The performance dataset includes the typical aircraft dimension variables, including fuselage height, length, width, wing dimensions, and empty/max takeoff weights. It also includes performance characteristics such as service ceiling, maximum range (also called ferry range), and maximum speed. While the dimension variables typically do not exhibit much variation between sources, the performance characteristics seem to be open to much greater interpretation. While we attempted to use Jane's as the sole source for all of the performance data, its incomplete reporting necessitate the use of Periscope and the Factsheets in some cases. Ultimately, performance data, particularly the ranges and speeds, can be reported in many different ways and should be cautiously employed in models. That being said, we collected the types of performance data most likely to remain consistent between sources.

The crew variable counts individuals that are part of the flight crew (pilot, co-pilot, etc.) and individuals that are part of the mission crew. For example, the MC-130 crew count would include the gun operators and the E-3A AWACS crew would include the radar operators also called specialists. The crew numbers do not include passengers.

Based on recommendations of model builders at the Pentagon, we also included indicators for afterburner and stealth. Both of these characteristics tend to correlate with significant increases in O&S costs, particularly stealth.

We included the logs of the dimension variables and the performance variables. Depending on the model diagnostics it may or may not make sense to transform these variables into log-space. However, many of said variables have wide variation between aircraft, so it’s not unreasonable to consider a transformation prior to running the models.

MD and MDS Aggregation

Most of the AFTOC and REMIS fields are additive in the sense that one can accumulate them at a higher level of aggregation by adding them together. However, categories such as average age cannot be averaged at the higher level, since it may distort that actual value. For the performance variables, the problem is similar to that of age—one must construct a weighted average of fields of similar and dissimilar airplanes within a single MDS or MD. Since the TAI variable exists in the dataset, we can create a weighted average for all of the weight, dimension and performance characteristics. However, depending on how the indicators aggregate (e.g. some with afterburner and some not within an MD), there may be further work. It turns out that at the MD level, there were no problems with the indicators—no within-MD differences between afterburner and stealth. However, to aggregate to the type level would probably involve abandoning the indicators or finding a better way to represent part of the fighter fleet as being stealth or having afterburners.

Models with Performance Variables

One of the important aspects to building a regression model with performance variables is that many of the variables are highly correlated and, therefore, will tend to explain variation in a similar way. The challenge here is to reduce the possibility of multicollinearity, without creating omitted variable bias. In other modeling circumstances, such as including highly correlated income and wealth variables, removing one may cause specification error. Whereas wealth and income are highly correlated, economic theory states that they account for two separate and distinct aspects of a person's financial status and should both be included in the model specification. However, most cost modeling does not include multiple aspects of dimension and performance, variables that are necessarily correlated due to the physics of flying. It is sufficient to include a subset of the performance variables to reduce multicollinearity and enhance model parsimony, while avoiding specification error.

Our original performance dataset included multiple highly correlated variables, with many pairs having a correlation of 0.9 or greater. For example, the correlation between Max Takeoff Weight and Empty Weight is 0.9157 in our dataset. Length, Height, Wingspan, and Wing Area are highly correlated also; we will retain Wingspan from these since its value is most consistent between data sources. Removing the sets of highly correlated variables allows for easier model building, since it mitigates much of the multicollinearity beforehand. This likely produces a better model specification, ultimately. Therefore, in addition to flying hours and Age, we will include the following performance variables: Engines, TotalThrust, Crew, Empty Weight, Ceiling, Maxspeed, Max Range, Stealth, and Afterburner.

We performed model selection, keeping in mind that we wanted this specification as close as the indicator as possible in order to compare them. There are other possible "performance" specifications, such as adding FH/Landing, that perhaps make better statistical models, with higher R^2. However, the specific goal of this model was to compare very closely with the MD indicator model—which did not include FH/landing. In order to more closely approximate the variation explained by the indicator variables, we chose not to transform the performance variables into log-space. However, the log-transformed performance model provides similar results for the flying hour coefficient.

Table A-3 shows the results for the performance specification for Total CAIG costs (w/o CAIG 2.1): $\ln(Cost_{my}) = \beta_0 + \beta_1 * Age_{my} + \beta_2 * \ln(Flying\ Hours_{my}) + \beta_3 * Engines_{my} + \beta_4 * Crew_{my} + \beta_5 * Ceiling_{my} + \beta_6 * Maxspeed_{my} + \beta_7 * Stealth_{my} + \beta_8 * Afterburner_{my} + \varepsilon_{my}$ where Engines, Crew, Ceiling and Maxspeed are continuous variables and Stealth and Afterburner are indicator variables.

Table A-3 Performance Specification Results

```
Linear regression                                      Number of obs =     361
                                                       F(  8,   352) =  248.00
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.8292
                                                       Root MSE      =  .60915
```

lncaigtot21	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf. Interval]	
lnflyinghrs	**.8317923**	**.0306339**	**27.15**	**0.000**	**.7715438**	**.8920409**
age	.0091597	.0032068	2.86	0.005	.0028529	.0154666
engines	.2532249	.0342628	7.39	0.000	.1858394	.3206104
crew	.0684921	.0085414	8.02	0.000	.0516935	.0852907
ceiling	.0000248	3.72e-06	6.68	0.000	.0000175	.0000322
maxspeed	.0010815	.0002159	5.01	0.000	.000657	.001506
stealth	1.258069	.0959987	13.11	0.000	1.069266	1.446872
afterburner	-.6368116	.1348052	-4.72	0.000	-.9019366	-.3716866
_cons	7.937324	.3170917	25.03	0.000	7.313691	8.560956

Table A-3 provides the results from the performance specification. The comparison of the coefficient on Ln(flying hours) between the baseline indicator specification (0.56) and the performance specification (0.83) shows that the performance specification may not be a very close proxy for the indicators. An interesting aspect to the performance model is that age has far less of an impact than it did in the indicator specification. Also, the performance specification has much narrowed confidence intervals on the independent variables, but there is a problem with the specification beyond it not being a decent proxy.

Figure A-2 Comparison of 95% CIs for Indicator and Performance Specifications

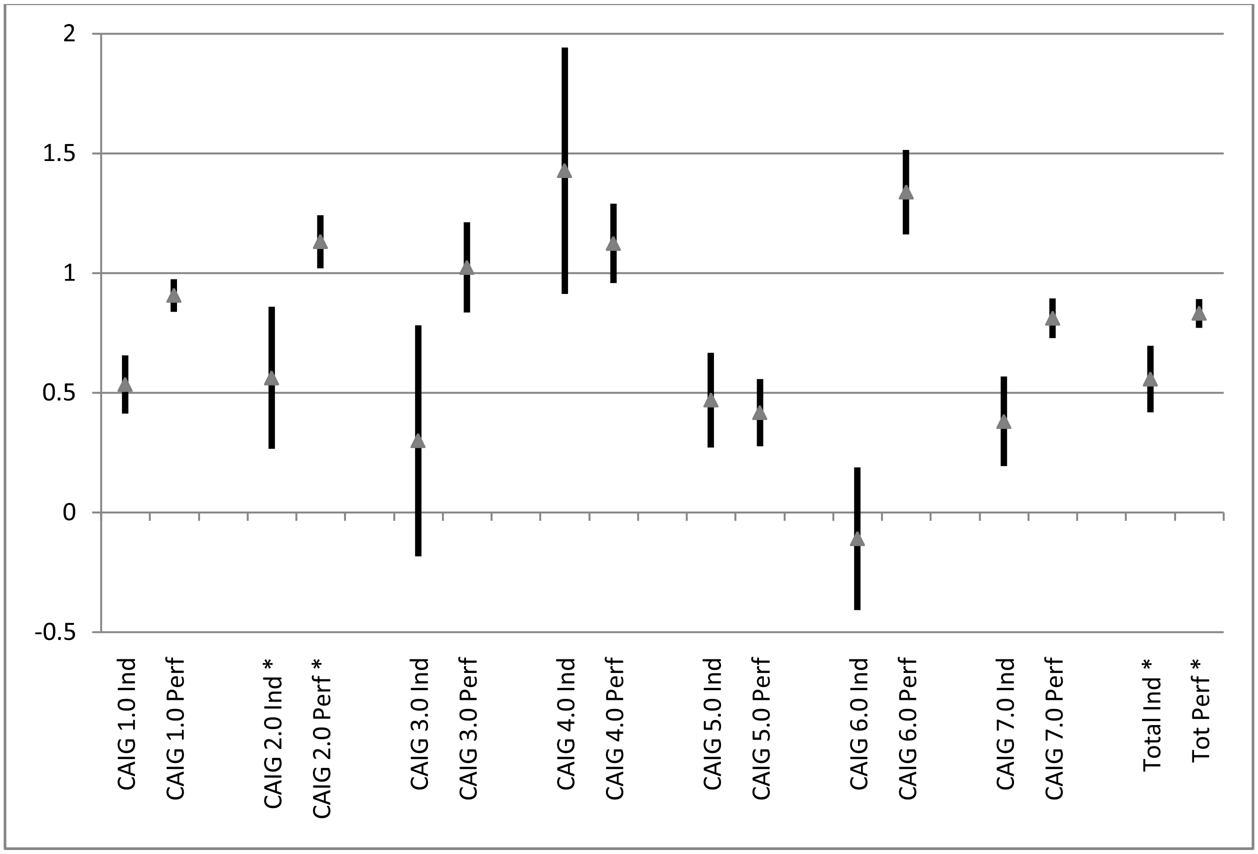


Figure A-2 shows that when comparing the indicator and performance specifications for the disaggregated CAIG level 1 costs—the level-one models—the confidence do not overlap in most cases. Notably, the two specifications align better on CAIG 4.0 and CAIG 5.0, but do not overlap in the other categories. Moreover, the specifications differ substantially on CAIG 2.0 costs, those most closely associated with the CPFH metric. The asterisk (*) in the table denotes that the CAIG 2.0 and CAIG Total costs do not include CAIG 2.1 fuel costs.

Performance Variables Versus Indicators Nested F-Test

Assertion: The Baseline Indicator specification and the Performance specifications are necessarily "nested" since the performance and dimension characteristics uniquely describe a particular MD/MDS, as do the indicators. In other words, each MD/MDS has a specific weight, range, thrust combination that separates it from all of the other MD/MDSs.

At the MD aggregation of costs, we test the following models:

Indicator Specification:
$\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + + \mu_m + \varepsilon_{my}$

Performance Specification:
$\ln(Cost_{my}) = \beta_0 + \beta_1 * Age_{my} + \beta_2 * \ln(Flying\ Hours_{my}) + \beta_3 * Engines_{my} + \beta_4 * Crew_{my} + \beta_5 * Ceiling_{my} + \beta_6 * Maxspeed_{my} + \beta_7 * Stealth_{my} + \beta_8 * Afterburner_{my} + \varepsilon_{my}$

F-test between Baseline Indicator model (results from A-2) and Performance model (results from A-3):

$$F_{(K_2-K_1)(N-K_2-1)} = \frac{(R_2^2 - R_1^2)/(K_2 - K_1)}{(1-R_2^2)/(N-K_2-1)}$$

```
F = [(0.9729 - 0.8292)/(36-9)] / [(1-0.9729)/(361-36-1)]
  = [(0.1437)/(27)] / [(0.0271)/(324)]
  = [0.005322 / [0.000084]
  = 63.63   (df = 27, 324)    F* = 1.52 (0.05)   F* = 1.80 (0.01)
```

So, the performance model does not perform as well as the indicator model, in that there is not statistical evidence in this test that shows we can replace the indicators with performance variables. Interestingly, the performance models have a narrower 95% CI than the Indicator specification, but a lower R^2. Despite the differences in the flying hour coefficient between the specifications, the results for the Performance specification support the basic conclusion that coefficient on Ln(Flying Hours) is less than one.

Flying Hour per Landing

Figure A-3 Flying Hour per Landing Trend FY96-FY06

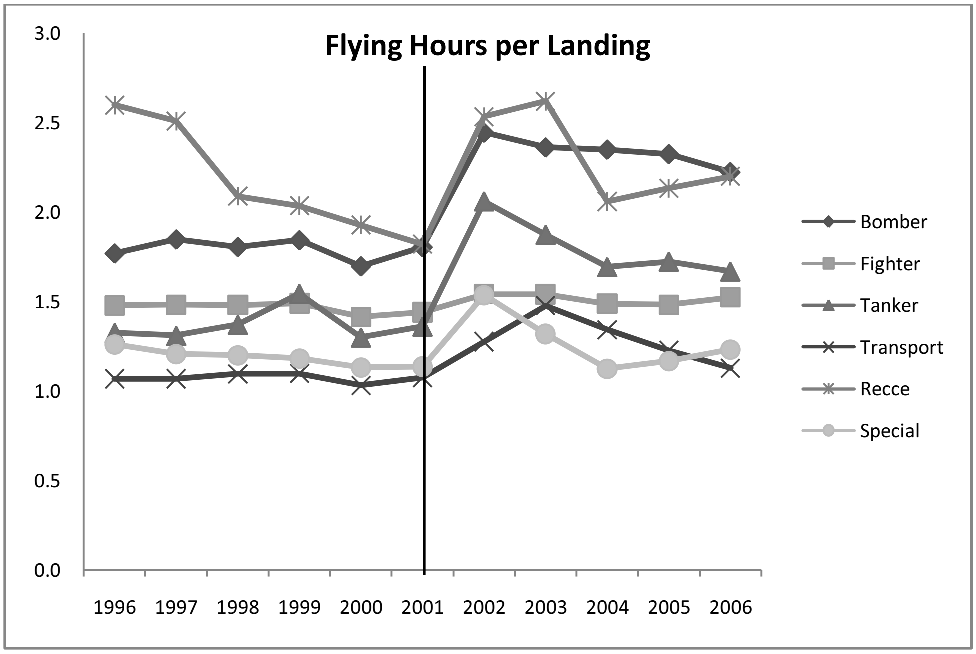


Figure A-3 shows the trends for FH/Landing for the seven basic aircraft types in this study. The chart for flying hours per landing shows a very similar pattern to the above chart. Models based on either of these two variables would produce very similar results.

Average Sortie Duration Regression Results

Table A-4 shows the results for the indicator specification with ASD variable for Total CAIG costs (w/o CAIG 2.1).

Table A-4 ASD Indicator Specification

Linear regression

Number of obs	=	361
F(36, 324)	=	1285.20
Prob > F	=	0.0000
R-squared	=	0.9732
Root MSE	=	.25173

lncaigtot21	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnflyinghrs	.5674055	.0753162	7.53	0.000	.4192351	.715576
age	.0592937	.0049281	12.03	0.000	.0495987	.0689888
lnfh_sortie	-.2570222	.2434934	-1.06	0.292	-.7360498	.2220054
a010	-.1306012	.0601022	-2.17	0.031	-.248841	-.0123614
ac130	.0014517	.3459981	0.00	0.997	-.6792347	.6821381
at038	-2.673548	.2299427	-11.63	0.000	-3.125917	-2.221178
b001	1.229774	.3104237	3.96	0.000	.6190731	1.840474
b002	1.638616	.4543335	3.61	0.000	.7448	2.532432
b052	-.5442689	.3656948	-1.49	0.138	-1.263705	.1751671
c005	.4305493	.2275671	1.89	0.059	-.0171464	.878245
c009	-1.51126	.1832879	-8.25	0.000	-1.871844	-1.150675
c017	.4739492	.271624	1.74	0.082	-.0604202	1.008319
c020	-1.034254	.29603	-3.49	0.001	-1.616638	-.4518703
c021	-1.681972	.1184993	-14.19	0.000	-1.915097	-1.448846
c026	-1.693598	.3211613	-5.27	0.000	-2.325423	-1.061773
c037	-.4973401	.3628967	-1.37	0.171	-1.211271	.2165912
c141	-.8506861	.1729161	-4.92	0.000	-1.190866	-.510506
e003	.3384506	.4024298	0.84	0.401	-.4532547	1.130156
e008	1.425831	.5350437	2.66	0.008	.3732332	2.47843
ec130	-2.22825	.3962923	-5.62	0.000	-3.007881	-1.448619
f015	.7939819	.091699	8.66	0.000	.6135812	.9743826
f016	.9319967	.1230453	7.57	0.000	.6899281	1.174065
f117	.2313039	.2112832	1.09	0.274	-.1843563	.646964
hc130	-1.295551	.2657397	-4.88	0.000	-1.818344	-.7727581
kc010	-.255334	.2841503	-0.90	0.370	-.8143464	.3036785
kc135	-.8153299	.1521422	-5.36	0.000	-1.114641	-.5160187
lc130	-.9871198	.3640283	-2.71	0.007	-1.703277	-.2709623
mc130	-.4862994	.1931868	-2.52	0.012	-.8663582	-.1062405
rc135	-.763197	.4332997	-1.76	0.079	-1.615633	.089239
t001	-1.195911	.121851	-9.81	0.000	-1.43563	-.956192
t006	-1.141968	.1834677	-6.22	0.000	-1.502906	-.7810294
t037	-2.531776	.1739157	-14.56	0.000	-2.873923	-2.18963
t038	-2.072721	.1773856	-11.68	0.000	-2.421694	-1.723748
t043	-2.501014	.3756988	-6.66	0.000	-3.240131	-1.761897
u002	.4651934	.3293495	1.41	0.159	-.1827402	1.113127
wc130	-2.177107	.392261	-5.55	0.000	-2.948807	-1.405407
_cons	13.27508	.8716555	15.23	0.000	11.56026	14.9899

Flying Hour/Total Active inventory (TAI) Regression Results

Table A-5 shows the results for the indicator specification with FH/TAI variable for Total CAIG costs (w/o CAIG 2.1).

Table A-5 FH/TAI indicator Specification

Linear regression

Number of obs	= 361
F(36, 324)	= 1445.34
Prob > F	= 0.0000
R-squared	= 0.9745
Root MSE	= .24542

lncaigtot21	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
lnflyinghrs	.6788335	.0895418	7.58	0.000	.5026766	.8549903
age	.0611358	.0050408	12.13	0.000	.051219	.0710526
lnfh_tai	-.4649065	.1741936	-2.67	0.008	-.8075999	-.1222131
a010	-.1531006	.0589672	-2.60	0.010	-.2691076	-.0370936
ac130	.2044534	.2954551	0.69	0.489	-.3767992	.7857059
at038	-2.502874	.189022	-13.24	0.000	-2.874739	-2.131008
b001	1.113495	.1787447	6.23	0.000	.7618481	1.465142
b002	1.642017	.317556	5.17	0.000	1.017285	2.266749
b052	-.8562904	.1584243	-5.41	0.000	-1.167961	-.5446203
c005	.4919299	.1448572	3.40	0.001	.2069504	.7769094
c009	-.8582357	.3423866	-2.51	0.013	-1.531817	-.1846543
c017	.9356879	.3280944	2.85	0.005	.2902236	1.581152
c020	-.5841522	.3598874	-1.62	0.106	-1.292163	.123859
c021	-1.294755	.203209	-6.37	0.000	-1.69453	-.894979
c026	-1.127688	.4526693	-2.49	0.013	-2.01823	-.237146
c037	.177251	.4628378	0.38	0.702	-.7332958	1.087798
c141	-.7663059	.1419876	-5.40	0.000	-1.04564	-.4869719
e003	.4808178	.2684813	1.79	0.074	-.0473689	1.009005
e008	1.735019	.4173887	4.16	0.000	.9138849	2.556153
ec130	-1.950846	.3609102	-5.41	0.000	-2.660869	-1.240823
f015	.6597136	.0936875	7.04	0.000	.475401	.8440262
f016	.7376143	.12957	5.69	0.000	.4827095	.9925191
f117	.3495441	.218964	1.60	0.111	-.0812265	.7803148
hc130	-1.165625	.2447963	-4.76	0.000	-1.647216	-.6840341
kc010	.0531631	.2571091	0.21	0.836	-.4526509	.5589771
kc135	-1.065242	.1036965	-10.27	0.000	-1.269245	-.8612381
lc130	-.4951734	.4181333	-1.18	0.237	-1.317772	.3274256
mc130	-.3502249	.1833408	-1.91	0.057	-.7109135	.0104638
rc135	-.5758113	.3077188	-1.87	0.062	-1.18119	.0295678
t001	-1.00145	.1565074	-6.40	0.000	-1.309349	-.6935512
t006	-.8045376	.2220289	-3.62	0.000	-1.241338	-.3677374
t037	-2.419018	.0907634	-26.65	0.000	-2.597577	-2.240458
t038	-2.085243	.1022641	-20.39	0.000	-2.286428	-1.884057
t043	-2.172908	.3615841	-6.01	0.000	-2.884257	-1.461559
u002	.680099	.284561	2.39	0.017	.1202785	1.239919
wc130	-2.049335	.3321827	-6.17	0.000	-2.702842	-1.395827
_cons	14.48547	.8041154	18.01	0.000	12.90352	16.06742

Results for CAIG Level-One Models

Tables A-5 and A-6 show the shows the results for the indicator specification with ASD and the indicator specification with FH/TAI variable, respectively, for Level-One CAIG costs.

Table A-6 Average Sortie Duration Specification Results

	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAIG 1.0	0.5585290	0.0646133	8.64	0.0000	0.4314129	0.6856450
CAIG 2.0 w/o 2.1	0.5958033	0.1655858	3.60	0.0000	0.2699333	0.9216733
CAIG 3.0	0.4292841	0.3018821	1.42	0.1590	-0.1710416	1.0296100
CAIG 4.0	1.9632170	0.2200880	8.92	0.0000	1.5295090	2.3969250
CAIG 5.0	0.4930163	0.1059005	4.66	0.0000	0.2846672	0.7013654
CAIG 6.0	0.0922855	0.1950377	0.47	0.6360	-0.2915401	0.4761112
CAIG 7.0	0.3819490	0.0980851	3.89	0.0000	0.1889849	0.5749131
Total w/o2.1	0.5674055	0.0753162	7.53	0.0000	0.4192351	0.7155760

Table A-7 FH/TAI Specification Results

	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
CAIG 1.0	0.6020123	0.1311334	4.59	0.0000	0.3440290	0.8599956
CAIG 2.0 w/o 2.1	0.7507625	0.1986440	3.78	0.0000	0.3598345	1.1416910
CAIG 3.0	0.4254675	0.2506420	1.70	0.0930	-0.0729617	0.9238967
CAIG 4.0	2.1402520	0.1656906	12.92	0.0000	1.8137410	2.4667640
CAIG 5.0	0.6769137	0.1262358	5.36	0.0000	0.4285569	0.9252706
CAIG 6.0	0.2813238	0.3198709	0.88	0.3800	-0.3481682	0.9108159
CAIG 7.0	0.4323874	0.1311298	3.30	0.0010	0.1744141	0.6903608
Total w/o2.1	0.6788335	0.0895418	7.58	0.0000	0.5026766	0.8549903

Chapter Five Regression Results

Table A-8 shows the results for the Flying Hour specification for each of the Level-Two CAIG elements:
FH specification: $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Flying\ Hours_{my}) + \mu_m + \varepsilon_{my}$

Table A-8 Flying Hour CAIG Specification Results

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncaig10	0.5342	0.0621	8.60	0.000	0.4120	0.6565
lncaig11	0.6312	0.0676	9.34	0.000	0.4982	0.7642
lncaig12	0.5704	0.0588	9.69	0.000	0.4546	0.6861
lncaig13	0.4462	0.0856	5.21	0.000	0.2778	0.6146
lncaig20no21	0.5615	0.1507	3.73	0.000	0.2649	0.8581
lncaig21	1.0890	0.0593	18.36	0.000	0.9723	1.2057
lncaig22	0.3620	0.1362	2.66	0.008	0.0941	0.6300
lncaig23	0.1042	0.1729	0.60	0.547	-0.2361	0.4445
lncaig24	0.4371	0.9397	0.47	0.643	-1.4287	2.3029
lncaig25	0.6786	0.1691	4.01	0.000	0.3459	1.0113
lncaig30	0.2996	0.2425	1.24	0.220	-0.1826	0.7817
lncaig40	1.9002	0.2229	8.53	0.000	1.4611	2.3393
lncaig41	1.4724	0.2854	5.16	0.000	0.9095	2.0354
lncaig43	1.4788	0.3145	4.70	0.000	0.8585	2.0991
lncaig44	0.7710	0.2490	3.10	0.002	0.2798	1.2621
lncaig50	0.4694	0.1006	4.67	0.000	0.2715	0.6673
lncaig52	0.5560	0.1146	4.85	0.000	0.3305	0.7815
lncaig53	0.3488	0.1834	1.90	0.059	-0.0128	0.7105
lncaig60	0.0243	0.1936	0.13	0.900	-0.3566	0.4052
lncaig61	0.0913	0.1642	0.56	0.579	-0.2319	0.4146
lncaig63	0.8669	0.1834	4.73	0.000	0.5051	1.2287
lncaig64	1.0794	0.1588	6.80	0.000	0.7666	1.3922
lncaig65	0.0294	0.2659	0.11	0.912	-0.4958	0.5547
lncaig66	-0.3531	1.2585	-0.28	0.780	-2.8618	2.1556
lncaig70	0.3795	0.0950	3.99	0.000	0.1926	0.5665
lncaig71	0.7230	0.1191	6.07	0.000	0.4887	0.9573
lncaig72	0.3597	0.0885	4.06	0.000	0.1856	0.5338
lncaigtot21	0.5567	0.0701	7.94	0.000	0.4187	0.6947

Table A-9 shows the results for the TAI specification for each of the Level-Two CAIG elements:
TAI specification: $\ln(Cost_{my}) = \alpha + \beta * Age_{my} + \delta * \ln(Total\ Aircraft\ Inventory_{my}) + \mu_m + \varepsilon_{my}$

Table A-9 TAI CAIG Specification Results

	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
lncaig10	0.6500	0.1123	5.79	0.000	0.4291	0.8710
lncaig11	0.7668	0.1209	6.35	0.000	0.5291	1.0046
lncaig12	0.7692	0.0769	10.00	0.000	0.6179	0.9204
lncaig13	0.6024	0.1080	5.58	0.000	0.3899	0.8149
lncaig20no21	0.7649	0.1978	3.87	0.000	0.3757	1.1542
lncaig21	1.2435	0.0776	16.03	0.000	1.0909	1.3962
lncaig22	0.5862	0.1851	3.17	0.002	0.2220	0.9504
lncaig23	0.2133	0.1886	1.13	0.259	-0.1580	0.5846
lncaig24	2.0267	1.0570	1.92	0.058	-0.0720	4.1254
lncaig25	0.9388	0.2222	4.23	0.000	0.5016	1.3760
lncaig30	0.4221	0.2467	1.71	0.091	-0.0685	0.9127
lncaig40	2.1391	0.1612	13.27	0.000	1.8215	2.4568
lncaig41	1.8310	0.2041	8.97	0.000	1.4284	2.2336
lncaig43	1.8035	0.4154	4.34	0.000	0.9842	2.6228
lncaig44	0.9474	0.2544	3.72	0.000	0.4457	1.4491
lncaig50	0.6638	0.1190	5.58	0.000	0.4297	0.8978
lncaig52	0.8110	0.1397	5.80	0.000	0.5362	1.0859
lncaig53	0.3940	0.2116	1.86	0.064	-0.0232	0.8111
lncaig60	0.2019	0.2950	0.68	0.494	-0.3786	0.7824
lncaig61	0.2471	0.2084	1.19	0.237	-0.1630	0.6571
lncaig63	1.1418	0.3104	3.68	0.000	0.5295	1.7542
lncaig64	1.2014	0.1517	7.92	0.000	0.9024	1.5003
lncaig65	-0.0846	0.2127	-0.40	0.692	-0.5048	0.3356
lncaig66	0.0746	1.3028	0.06	0.954	-2.5225	2.6717
lncaig70	0.4638	0.1266	3.66	0.000	0.2147	0.7129
lncaig71	0.9443	0.1372	6.88	0.000	0.6743	1.2142
lncaig72	0.4255	0.1198	3.55	0.000	0.1898	0.6612
lncaigtot21	0.7089	0.0859	8.25	0.000	0.5399	0.8778

Limitations

Limitations of MD Aggregated models

A Logistics Management Institute (LMI) study on the relationship between age and CPFH (Herzberg, et al, 2003) notes that there are several limitations to analyzing system-level aggregated data, in addition to dearth of observations. Summary level data tend to mask changes within individual systems such as accounting and maintenance system changes. Also, the averaging of variables can bias results. An example of this type of bias is acquiring new aircraft for an existing MD. New systems and older systems tend to have higher costs, according to the bathtub curve of maintenance costs. Adding new aircraft may have the effect of lowering the average age and increasing the costs.

One of the aspects of using log-log specifications is that it is relatively difficult to transform back to levels; but this is not necessary.[37] The goal of our analysis is not to build MD-specific cost factors, but rather to inform the functional form of the CPFH metric. To build better CPFH factors specific to an MD or MDS, one would likely use a level-level model for a given MD/MDS and incorporate detailed information unique to the MD/MDS.

Limitations of disaggregated AFCAIG cost categories

There is a possible problem with running models at the MDS level against the 28 disaggregated AFCAIG cost categories. Not every MDS within each MD shows costs in the same categories. Usually there are small differences—one or two categories—but we have not evaluated this in terms of FHs of costs. The primary problem is that unless we align the proper FHs with the costs, the MD specification will contain too many FHs in some categories. Ultimately, one might collect flying hours for all of the MDSs in a given MD, but only have costs for a subset of the MDSs represented.

[37] The log-log specification implies constant elasticity interpretation we are fundamentally interested in coefficient on Ln(FH) ceteris paribus. The model is multiplicative in level-level and additive in log-log.

References

Data Systems

Air Force Total Ownership Costs (AFTOC) data System. Accessible through the Air Force Portal, with permission from the AFTOC data manager.

Reliability & Maintainability Management Information System (REMIS) Data System. Accessible through the Air Force Portal, with permission from the data manager.

Air Force Statistical Digest. Provides description of Air Force resources, organizations, and activities including aircraft average ages, flying hours, and inventory. Accessible through the Air Force Portal, with permission from SAF/FM.

Bibliography

Air Force Cost Analysis Improvement Group (AFCAIG), Cost per Flying Hour Process Guide, Nov 1999.

Air Force Magazine, Air Force Almanac, Available at http://www.afa.org/magazine/almanacs.asp, Jan 2007.

Armstrong, Patrick D., Developing an Aggregate marginal Cost Per Flying hour Model for Air Force's F-15 Fighter Aircraft, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB OH, Mar 2006.

Bureau of Economic Analysis, Gross Domestic Product, Available at http://www.bea.gov, Jun 2007.

Bureau of Labor Statistics (BLS), Consumer Price Index—All Urban Consumers, Available at http://www.bls.gov/cpi/home.htm, May 2007.

Cobaniau, Aurel and Konrad Madej, Analysis and Forecasting of Operating and Support Costs for F-16C/D, MS Thesis, Naval Post Graduate School, Monterey, CA, Jun 2006.

Cooper, Major Dane, Air Force Cost Analysis Agency (AFCAA), Crystal City, Email correspondence, May 2007.

Department of the Air Force, Fact Sheets, Available at http://www.af.mil/factsheets, May 2007.

Department of Defense (DoD), Cost Analysis Improvement Group(CAIG), DoD Directive 5000.04, Available at http://www.dtic.mil/whs/directives/corres/pdf/500004p.pdf, Jun 2007.

Dixon, Matthew, The Maintenance Costs of Aging Aircraft: Insights from Commercial Aviation, Santa Monica, CA: RAND MG-486-AF, 2006.

Foster, N. W., and H. D. Hunsaker, "The Effect of Aircraft Age on Flying Hours and Maintenance Costs," Air Force Logistics Command Technical Report 82-099, Jun 1983.

General Accounting Office (GAO), “Observations on the Air force Flying hour Program”, Report to Congressional Committees, Washington, DC, GAO/NSIAD-99-165, Jul 1999.

Greenfield, Victoria A. and David M. Persselin, An Economic Framework for Evaluating Military Aircraft Replacement, Santa Monica, CA: RAND MR-1489-AF, 2002.

Guo, M. and others. “Effect of Environmental Factors on the Corrosion of 2024T3 Aluminum Alloy,” Materials Forum, Volume 28:433-438 (2004).

Hawkes, Eric M., Predicting the Cost per Flying Hour for the F-16 using Programmatic and Operational Variables, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB OH, Jun 2005.

Herzberg, Eric F., Joseph Callender II, David M. Cutter, Owen R. Thompson, and Andrew Timko, The Relationship Among Cost, Age, and Usage of Weapon Systems, Logistics Management Institute Report LG102T2, Jan 2003.

Hildebrandt, Gregory G. and Man-bing Sze, An Estimation of USAF Aircraft Operating and Support Cost Relations, Santa Monica, CA: RAND N-3062-ACQ, May, 1990.

Keating, Edward G., and Matthew Dixon, Investigating Optimal Replacement of Aging Air Force Systems, Santa Monica, CA: RAND MR-1763-AF, 2003.

Keating, Edward G., Don Snyder, Matthew Dixon and Elvira N. Loredo, Aging Aircraft Repair-Replacement Decisions with Depot-Level Capacity as a Policy Choice Variable, Santa Monica, CA: RAND MG-241-AF, 2005.

Laubacher, Matthew E., Analysis and Forecasting of Air Force Operating and support Cost for Rotary Aircraft, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB OH, Mar 2004.

Lies, Tom and Larry Klapper, “Air Force Cost and Performance (AFCAP) Tool,” Briefing presented at AFTOC Users Conference, Mar 2007.

Lively , Crash, AFTOC System expert, Air Force Cost Analysis Agency (AFCAA), Crystal City, Telephone discussion, May 2007

Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining, Introduction to Linear Regression Analysis, 3rd ed., New York, NY, Wiley and Sons, Inc., 2001.

Office of the Secretary of Defense Cost Analysis Improvement Group (OSD CAIG), “Operating and Support Cost-Estimating Guide,” Washington, DC, May 1992, Available at http://www.dtic.mil/pae, Jul 2007.

Pyles, Raymond A., Aging Aircraft: USAF Workload Consumption and Material Consumption Life Cycle Patterns, Santa Monica, CA: RAND MR-1641, 2003.

Rose, Pat A. Jr., “Cost Per Flying Hour Factors: A Background and Perspective of How They Are Developed and What They Do,” The Air Force Comptroller, 31-1:4-9, Jul 1997.

SAF/FMC, United States Air Force Statistical Digest, Fiscal Year 2005, Available through the Air Force Portal, www.my.af.mil, May 2007.

SAF/FMC, Logistics Cost Factors Description , Air Force Instruction 65-503, November 2005, Available at https://www.my.af.mil/gcss-af/afp40/USAF/ep/contentView.do?contentType=EDITORIAL&contentId=730346&programId=730401&pageId=681744&channelPageId=-1073755509, Jul 2007.

Sherbrooke, Craig C., Using Sorties vs. Flying Hours to Predict Aircraft Spares Demand, Logistics Management Institute Report AF501LN1, Apr 1997.

Slay, F. Michael, Demand Forecasting, Logistics Management Institute Report AF401N2, Aug 1995.

Wallace, John M., Scott A. Houser, and David A. Lee, A Physics Based Alternative to Cost-Per-Flying-Hour Models of Aircraft Consumption Costs, Logistics Management Institute Report AF909T1, Aug 2000.

Wooldridge, Jeffrey M., Introductory Econometrics: A Modern Approach, 2nd ed., Mason, Ohio, South-Western, 2003.